NASA CONTRACTOR
REPORT

NASA CR-61081

RESEARCH STUDIES OF RANDOM PROCESS THEORY AND PHYSICAL APPLICATIONS

Prepared Under Contract NAS8-11346 by

George W. Bordner, Charles J. Greaves, and Walter W. Wierwille

CORNELL AERONAUTICAL LABORATORY, INC.
Buffalo, New York

For

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER
Huntsville, Alabama

August 4, 1965

NASA CR-61081

RESEARCH STUDIES OF RANDOM PROCESS THEORY AND PHYSICAL APPLICATIONS

By

George W. Bordner, Charles J. Greaves, and Walter W. Wierwille

Prepared Under Contract NAS8-11346 by
CORNELL AERONAUTICAL LABORATORY, INC.
Buffalo, New York

For Computation Laboratory

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER

CONTENTS

CONTENTS (Cont.)

CONTENTS **(Cont.)**

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont'd)

ACKNOWLEDGEMENTS

On 1 July 1964, a project sponsored by the Computation Laboratory of the Marshall Space Flight Center, Huntsville, Alabama, was initiated with Cornel1 Aeronautical Laboratory to perform a research study of random process theory. The work was performed at CAL mainly in the Avionics Department. The Project Engineer was G. W. Bordner with Senior Investigators C. J. Greaves and Dr. W. W. Wierwille. Investigators were Dr. T. R. Benedict, W. D. Fryer, and G. A. Gagne.

The work described herein was supported in its entirety by the National Aeronautics and Space Administration, Huntsville, Alabama, under Contract No. NAS 8-11346. The contract technical monitor was Mr. Jack Jones, of the Computer Laboratory. The authors wish to express their appreciation to Mr. Jones and his colleague, Howard Newberry, for their helpful participation in this project.

RESEARCH STUDIES OF RANDOM PROCESS THEORY AND PHYSICAL APPLICATIONS

By George W. Bordner, Charles J. Greaves, and Walter W. Wierwille

CORNELL AERONAUTICAL LABORATORY, INC.
of Cornell University

SUMMARY

Many techniques of processing of random data have been presented which can be applied to engineering problems. Techniques for estimating correlation functions and power spectral densities for all classes of random processes, including nonstationary, stationary, gaussian bivariate, etc., in addition to deterministic processes have been described in a rigorous mathematical sense.

Further aids to the processing techniques are also given. These include a complete thesis on the synthesis of optimal digital filters, including low-pass, high-pass, band-pass, and notch filters. The technique of digital heterodyning is also presented. To further aid the processing engineer methods of applying the techniques on a general purpose digital computer are presented.

For the previously unproven processing techniques presented, experimental verification is given.

INTRODUCTION

On 1 July 1964 a project sponsored by the Computation Laboratory of the Marshall Space Flight Center, Huntsville, Alabama, was initiated with Cornel1 Aeronautical Laboratory (CAL) to perform a research study of random process theory. CAL was required to examine the existing Computation Laboratory techniques used to reduce and .analyze random process data toward the objective of devising new or improved applications of statistics and random process theory to existing effprt of endeavor. The random processes of interest were of meteorology, vibration, and acoustic measurements or any other applicable random process. Of particular concern was the physical application of many reduction processes and functions that were not yet commonly applied to the existing engineering problems at Marshall Space Flight Center.

The CAL study was designed to meet the computation Laboratory objectives. The specific goals of the research study were to reduce the data editing and computer usage time, to increase the "accuracy" of the statistical estimates of the processed data, and to recommend future applications of existing data reduction equipment. These improvements were to be a result of CAL's investigation of the techniques used at the Computation Laboratory and the appropriate application of:

1. Digital filtering techniques
2. Correlation function analysis
3. Spectral smoothing techniques
4. Special functions or processing
5. Spectrum analysis of nonstationary signals

Sponsor approval of these technical areas of work and a cross check with the requirements indicated a research program to meet the Computation Laboratory requirements. The above list of 5 technical areas of work were performed and the results given in the technical sections of this report. The following is an outline of the material presented.

A. Digital Filtering

- a) Time and frequency relationships
- b) Synthesis of optimum digital filters
- c) Optimum filters with the constraint of unity D. C. gain
- d) Notch filters
- e) Digital heterodyning

B. Correlation Function

a) Summary of available computational methods
b) "Half-polarity" correlator analysis
c) "Half-polarity". and "Full-precision" correlators compared
d) Correlation of signal plus noise
e) Computer techniques of calculating correlation functions

C. Optimal Spectral Smoothing

D. Deterministic Data Processing

E. Nonstationary Spectrum Analysis

a) Theory
b) Experimental results

F. Nonstationary Correlation Analysis

a) Theory
b) Experimental results

Conclusions to the above technical areas are given in the Conclusion Section of this report. One can always broaden the area of study for the type of research performed for this project. Further, one may perform more extensive research in the areas already discussed. Extensions as these are summarized in the section on Recommendations for Further Work.

DIGITAL FILTERING

The results of discrete-data processing is greatly dependent upon the selection of an appropriate sampling interval during the analog to digital conversion. [4]* In practice we know that no signal of finite duration has a band-limited spectrum. Under such circumstances the selection of the sampling interval must be based upon reasonable knowledge that the amplitude of the spectrum is negligible beyond some frequency fc . This particular frequency may well be beyond the highest frequency of interest however, the Nyquist sampling criteria must be satisfied at least to a good approximation, thus the sampling frequency to be selected is 2 fc.

Time-Frequency Relationships

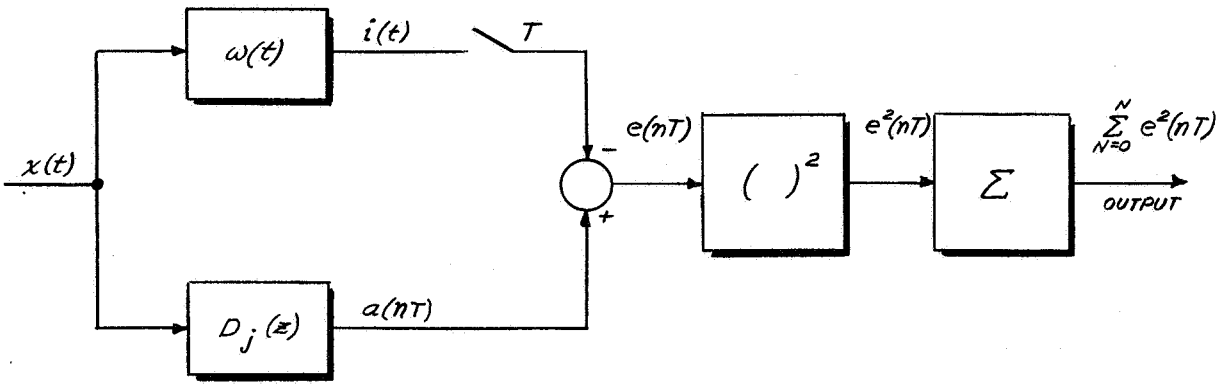
Where digital reduction and processing of data is employed, it is often desirable to perform various pre-processing filtering operations. Whitening and frequency band limiting are only two examples of pre-processing operations.

Extensive literature exists describing the many methods which may be employed in the simulation of transfer functions on general purpose digital computers. [11, 12, 13, 14, 15] It would be useful to relate the information available on time domain aspects of digitally simulated transfer functions and the frequency characteristics of these transfer functions. Consider for example that a given frequency characteristics is needed for spectral prewhitening. With the use of conventional techniques, such as a Bode plot, this given gain frequency information can be represented by a transfer function. This transfer function in turn is converted by techniques such as those cited in Reference 15 into a digital filter.

If from the many techniques one is chosen which yields the sum of the error squared in the time domain a minimum, it will also insure minimum power in the error signal. Equivalently, it minimizes the integral of the square of the difference between the Fourier transforms of the actual and ideal response.

In the following diagram $D_j(z)$, the method selected to digitally represent the transfer function whose impulse response is $\omega(t)$, is assumed to yield a minimum at the output.

*Superscripts refer to the references.



$$\text{define } \Phi_{ee}(kT) = \frac{1}{N}\sum_{n=0}^{N} e(nT)\,e(nT+kT) \qquad (1)$$

$$\text{Thus } \Phi_{ee}(0) = \frac{1}{N}\sum_{n=0}^{N} e^2(nT) \qquad (2)$$

Let $F_{ee}(z)$ be the sequence (or z) transform of $\phi_{ee}(kT)$ taking the inverse sequence transform we have

$$\phi_{ee}(kT) = \frac{1}{j2\pi}\oint_{\Gamma} F_{ee}(z)\,z^{k-1}\,dz \qquad (3)$$

where the path Γ is the unit circle. Evaluating both sides at $k = o$ gives

$$\phi_{ee}(0) = \frac{1}{N}\sum_{n=0}^{N} e^2(nT) = \frac{1}{j2\pi}\oint_{\Gamma} F_{ee}(z)\,\frac{dz}{z} \qquad (4)$$

changing from a contour integral by letting $z = e^{j\omega T}$

$$\frac{1}{N}\sum_{n=0}^{N} e^2(nT) = \frac{T}{2\pi}\int_{-\frac{\omega_s}{2}}^{\frac{\omega_s}{2}} \Phi_{ee}(\omega)\,d\omega \tag{5}$$

where $\omega_s = \frac{2\pi}{T}$

$$F_{ee}(e^{j\omega T}) = \Phi_{ee}(\omega) = \frac{|A(\omega)-I(\omega)|^2}{T} \tag{6}$$

The above equation, is essentially the discrete form of Parseval's Theorem, illustrating the desired results; namely, minimizing the left-hand side of equation (5) is equivalent to minimizing the right-hand side.

In general, then, the method of digitally simulating an analog filter which provides a best time response (with the mean-square-error criteria) is also the one which provides the least distortion from the ideal frequency response.

Double-Ended Filters

In recent years a considerable amount of literature has been written on digital filtering techniques. Unfortunately, much of the present literature does not explicitly state the basis of the synthesis. Furthermore, in many cases a very general class of filters is considered, thereby making the resulting expression for the filter weights somewhat cumbersome to apply.

The class of filters to be considered here is limited to those most useful to the pre-processing of data from which correlation and spectral density functions will be obtained.

Optimality of the weighting sequences is based upon the minimization of the following error index

$$J = \int_{-\omega_c}^{\omega_c} \left| F_a(\omega) - F_i(\omega) \right|^2 d\omega \tag{7}$$

where $F_a(\omega)$ is the actual frequency function achieved and $F_i(\omega)$ is the desired ideal frequency function. Details of the minimization procedure are carried out in Appendix A.

In nonreal time processing of discrete-data it is feasible and often advantageous to consider the use of double-ended filters. Such a filter would respond before the input arrived and the weighting sequence would have values for both positive and negative time hence the term double-ended.

The class of filters to be investigated is defined as the set of all filters which can be expressed as a linear combination of filters of the following form:

$$F_\ell(\omega) = R_\ell(\omega) + j I_\ell(\omega) \tag{8}$$

where $R_\ell(\omega)$, $I_\ell(\omega)$ are real functions of the real frequency variable

and

$$I_\ell(\omega) = 0$$
$$R_\ell(\omega) = \begin{cases} \\ \end{cases} \quad ; \; |\omega| \leq \omega_\ell \leq \omega_c \quad \text{otherwise} \tag{9}$$

That is to say $F(\omega)$ is a member of the class if $F(\omega)$ can be expressed in the following form:

$$F(\omega) = \sum_{\ell=0}^{i} k_\ell F_\ell(\omega) \tag{10}$$

where k_ℓ is any arbitrary constant.

The relationship between the folding frequency and spacing between data points is given by

$$T = \frac{1}{2 f_c} = \frac{\pi}{\omega_c} \tag{11}$$

where

T = time spacing between data points

$\omega_c = 2\pi f_c$

Two properties of this class of filters are apparent from the definition given in equation (9). $I(\omega)$ being zero implies that the phase function is zero at all frequencies and thus no phase distortion of the input signal is produced. Secondly, the impulse response is symmetrical about the origin as is readily shown by taking the inverse Fourier transform of $F(\omega)$.

Consider now the synthesis of a low-pass discrete-data filter whose ideal frequency characteristics are shown in Figure 1 .

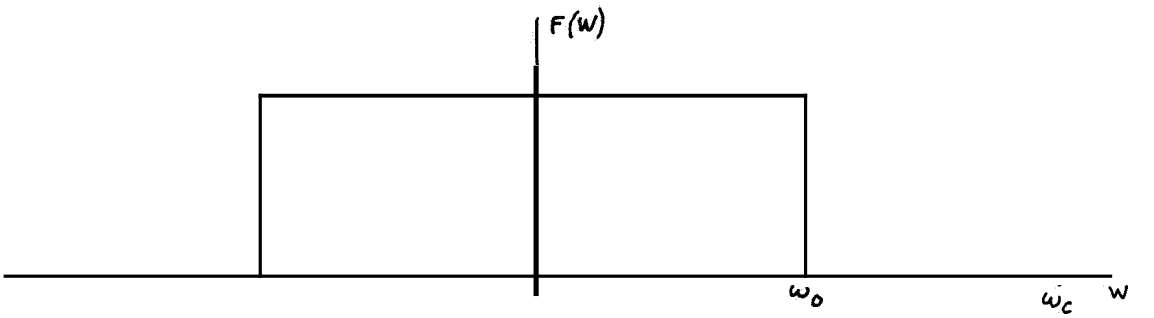


Figure 1 IDEAL FREQUENCY CHARACTERISTICS

Taking the inverse tranform of $F(\omega)$ yields

$$\mathcal{F}\left[F(\omega)\right] \quad \frac{1}{2\pi}\int e^{j\omega t} \tag{12}$$

$$f(t) = \frac{\omega_0}{\pi}\,\frac{\sin \omega_0 t}{\omega_0 t}$$

The signal $f(t)$ is frequency band-limited and thus can be sampled every T seconds without causing frequency folding. The function after having been sampled and multiplied by T to restore the frequency gain is given by

$$f_n = \frac{\omega_0}{\omega_c}\,\frac{\sin n\frac{\omega_0}{\omega_c}\pi}{n\frac{\omega_0}{\omega_c}\pi} \tag{13}$$

Frequency characteristics of the sequence f_n is illustrated in Figure 2, where only two of the infinite number of shifted components have been included.

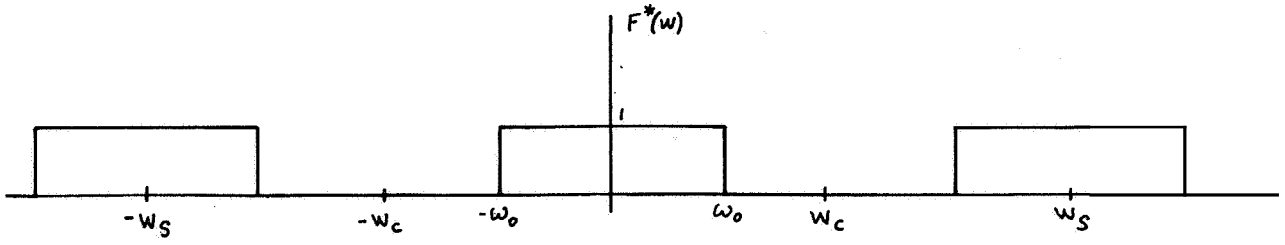


Figure 2 FREQUENCY CHARACTERISTICS OF f_n

If the input data is designated by x_n then the output of this low-pass filter y_n is given by the discrete convolution.

$$y_n = x_n * f_n = \sum_{k=-\infty}^{\infty} x_{n-k} f_k \qquad (14)$$

From the basic low-pass discrete-data filter described above generation of a few of the more commonly used filters is considered. A high-pass filter is formed by the substraction of a low-pass from an all pass or unity as shown in Figure 3 .

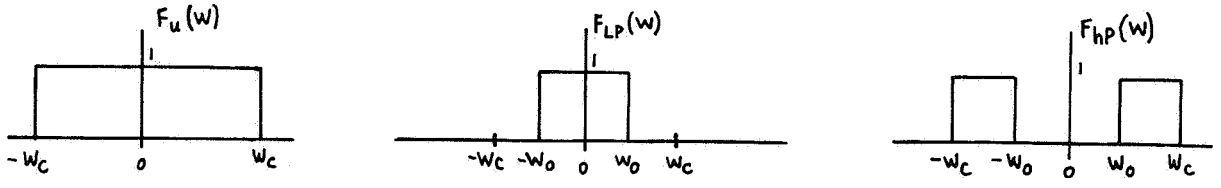


Figure 3 FILTER COMBINATIONS

The weighting sequence for the high-pass filter is given by the sampled inverse Fourier Transform of

$$F_u(\omega) - F_{LP}(\omega)$$

which yields

$$f_{hp}(n) = 1 - f_{LP}(n) \tag{15}$$

through similar reasoning the generation of the weighting sequence for a band-pass filter would be given by

$$f_{bp}(n) = f_{LP2}(n) - f_{LP1}(n) \tag{16}$$

Finally the notch filter is arrived at from the substraction of a band-pass from an all pass and the weighting sequence for the notch is

$$f_N(n) = 1 - f_{bp}(n) \tag{17}$$

A variety of other freqeuncy functions can readily be achieved through various combinations of the basic low-pass forms represented by equation (9).

The weighting sequence arrived at in the foregoing discussion spans the entire time axis. From the standpoint of practical considerations, the sequence may only extend over some finite interval. Since the class of filters which have been considered are derived from the basic low-pass, the frequency characteristics of this filter using only a finite length of its weighting sequence is investigated.

The frequency characteristics of such a truncated sequence is given by

$$F^*(\omega) = \sum_{n=-N}^{N} f_n e^{-jn\omega T} \tag{18}$$

utilizing the property that $f_{-n} = f_n$ the above expression can be placed in the following form

$$F^*(\omega) = f_0 + 2\sum_{n=1}^{N} f_n \cos n\frac{\omega}{\omega_c}\pi \tag{19}$$

if the following changes of variables are made

$$\frac{\omega}{\omega_c} = \alpha \quad ; \quad \frac{\omega_0}{\omega_c} = \beta$$

the normalized frequency function is given by

$$\bar{F}^*(\alpha) = \beta\left(1 + 2\sum_{n=1}^{N} \frac{\sin n\beta\pi}{n\beta\pi}\cos n\alpha\pi\right) \tag{20}$$

A series of normalized frequency plots is included in Appendix B. The relationships between sharpness of cutoff and the parameter N is

readily determined from an inspection of these graphs. It is felt that the generation of these plots, which are not commonly available, would in general be of greater aid in the determination of filter weights than a relationship between the value of the error index and N .

Often it is desirable for a given numerical filter to exhibit unity D. C. gain in addition to other specified characteristics, To obtain a solution for the optimum filter weights, which passes D. C. with unity gain, the constraint is multiplied by λ a positive constant and added to the original performance index. The filter weights obtained with this approach are given by

$$c_n = \sum_{\ell=0}^{L} k_\ell f_{\ell n} + \frac{1 - \sum_{j=-N}^{N} \sum_{\ell=0}^{L} k_\ell f_{\ell j}}{2N+1} \qquad (21)$$

Equation (21) gives the optimum weights in terms of the component filter weights. Details of the derivation of equation (21) are found in Appendix C .

It is interesting to note that the filter weights specified by equation (21) differ from those of Reference **22** which also exhibit unity D. C. gain. The discrepancy between the **two** results apparently arises from the fact that no attempt was made in Reference 22 to obtain the weights based upon the minimization of a performance index.

Notch Filters

In the calculation of power spectra, the effect of frequency spreading due to the use of any particular spectral window is minimized when the spectrum under consideration is prewhitened before final spectral analysis is performed. The shape and location of the prewhitening or notch filters would be determined by a fast pilot estimation. Computational time for spectral analysis may increase in general with the implementation of notch filters, along with their inverses, to restore the spectral estimate. However, the additional time taken, if any, must be considered as a trade off for greater accuracy in the spectral estimates.

A portion of the prewhitening may be achieved by a modification of the analog transducer transmission system; however, in almost all cases, greater flexibility is provided by complete digital processing.

As an alternative to the double-ended discrete filter technique, consider the following bilinear transformation given by:

$$S = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \qquad (22)$$

which maps the entire left half of the complex S plane within the unit circle in the complex z plane. The application of this transformation upon linear constant parameter analog filters to produce a discrete-data recursive formula is often referred to as the Tustin transformation. Simplicity and speed with which the recursive weights are computed make this technique advantageous. In general terms, consider the filter given by

$$F(S) = \frac{\sum_{j=0}^{\beta} a_j S^j}{\sum_{j=0}^{\beta} b_j S^j} \tag{23}$$

Applying the Tustin transformation to equation (23) yields the following recursive formula

$$y_m = \sum_{i=0}^{\beta} \frac{A_i}{B_0} x_{m-i} - \sum_{i=0}^{\beta} \frac{B_i}{B_0} y_{m-i} \tag{24}$$

where the weights are determined by

$$R_i = \sum_{j=0}^{\beta} r_j 2^j T^{\beta-j} \sum_{l=0}^{i} (-1)^l \binom{j}{l} \binom{\beta-j}{i-l} \tag{25}$$

where

$R_j = A_j$ or B_j

and $r_j = a_j$ or b_j respectively.

As a specific example, a notch filter with the analog pass characteristics illustrated in Figure 4 is given below:

$$\frac{S^2 + \omega_o^2}{S^2 + 2\omega_o S + \omega_c^2} \tag{26}$$

Application of equation (25) yields the following weights:

$$\begin{aligned} A_0 &= \omega_0^2 T^2 + 4 \\ A_1 &= 2\omega_0^2 T^2 - 8 \\ A_2 &= \omega_0^2 T^2 + 4 \\ B_0 &= \omega_0^2 T^2 + 4\omega_0 T + 4 \\ B_1 &= 2\omega_0^2 T^2 - 8 \\ B_2 &= \omega_0^2 T^2 - 4\omega_0 T + 4 \end{aligned} \tag{27}$$

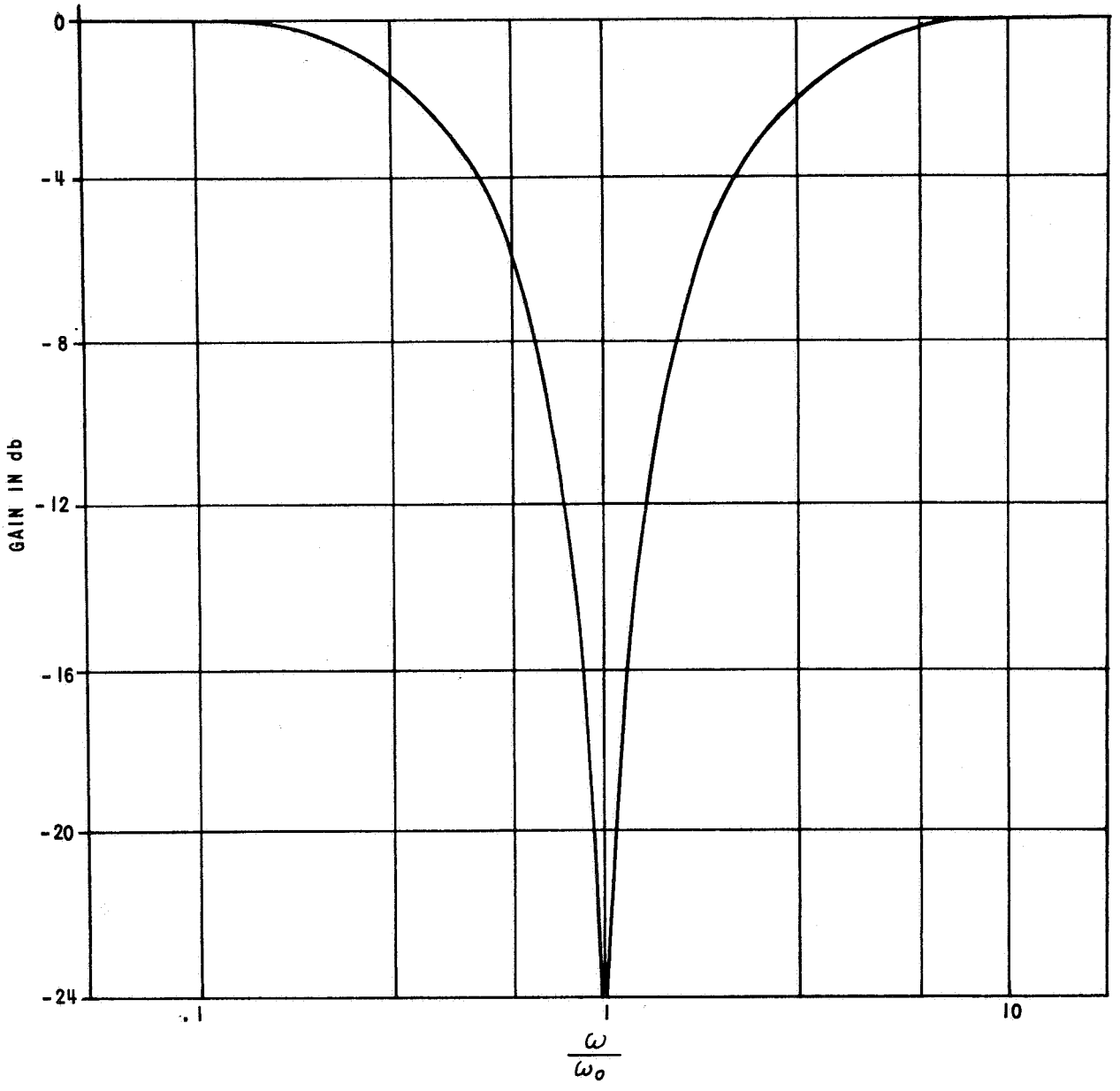


Figure 4 GAIN vs NORMALIZED FREQUENCY

The discrete-data frequency characteristics will be similar to that of Figure 4 when $\omega_0 T << 2$. If this condition is not satisfied, the following relationship should be used:

$$\omega_0 = \frac{2}{T}\sqrt{\frac{1-\cos\omega_1 T}{1+\cos\omega_1 T}} \tag{28}$$

where ω_1 is the rejection frequency desired in the discrete filter and ω_0 is the value used in equation (26) before the Tustin transformation is made. When total frequency rejection is not desired, damping is added to the numerator of equation (26).

Heterodyning

In situations where the power spectrum is desired for only a band of frequencies rather than the complete response, an appreciable savings of computational time may be achieved by filtering and heterodyning before spectral analysis is performed. The actual time saving depends upon the width of the band of interest, and its location within the spectrum. The advantage of this technique will diminish as the band of interest approaches the origin (i.e., zero frequency).

Before taking a quantitative look at the trade-offs involved in this heterodyning scheme, consider the system diagram and the characteristics of the specific filters. The following discussion of the filtering-heterodyning process concerns discrete data; however, similar arguments can be stated for analog data reduction techniques.

The processing to be performed is illustrated in Figure 5. The output of the filter heterodyne system $\mathcal{A}_3(nT)$ is given in the time domain by

$$s_3(nT) = \left\{\left[s_0(nT) * f_1(nT)\right] \cdot h(nT)\right\} * f_2(nT) \tag{29}$$

Where, by definition

$$f_a(nT) * f_b(nT) = \sum_{k=-\infty}^{\infty} f_a(kT)\, f_b(nT-kT) \tag{30}$$

is the discrete convolution. Taking the Fourier transform of equation (29) yields the following expression for the frequency response of $\mathcal{A}_3(nT)$.

$$S_3(\omega) = \left\{\left[S_0(\omega) \cdot F_1(\omega)\right] * H(\omega)\right\} \cdot F_2(\omega) \tag{31}$$

The signal $\mathcal{A}_3(nT)$ is processed in the usual manner to obtain its autocorrelation function and power spectrum.

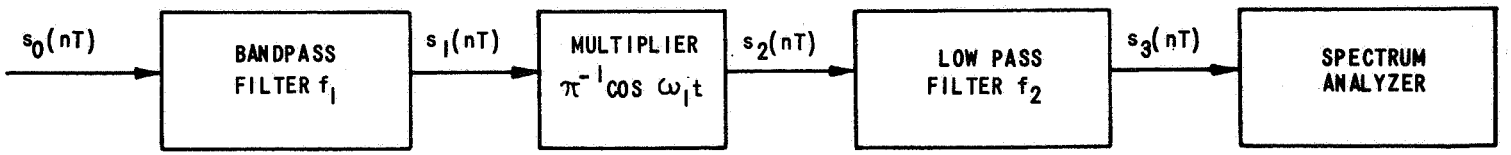


Figure 5 DIGITAL HETERODYNER

To demonstrate the effect of this signal processor, an example is given. Shown in Figure 6 is a sample power spectrum $P_o(\omega)$ of $A_i(nT)$ and three ideal filter frequency responses along with the power spectrum $P_3(\omega)$ of $A_3(nT)$.

Considering a typical numerical example, let

$$\begin{aligned} f_c &= 3500 \text{ cps} \\ w_2 &= .8 w_c \\ w_1 &= .5 w_c \\ \Delta w &= w_2 - w_1 = .3 w_c \\ T &= 1/7000 \end{aligned} \tag{32}$$

length of record = 1 second

The double-ended filter responses are given by

$$f_1(n) = .3 \cos n .65\pi \quad \frac{\sin n .15\pi}{n .15\pi} \tag{33}$$

$$f_2(n) = \frac{.3 w_c}{\pi} \quad \frac{\sin n .3\pi}{n .3\pi} \tag{34}$$

Since the band from w_1 to w_2 has been heterodyned down to 0 to Δw, a new sampling interval, $T_{\Delta w}$, can be used for processing $A_3(nT)$. Then,

$$\frac{T_{\Delta w}}{T} = \frac{w_c}{\Delta w} = \frac{10}{3}$$

Truncating to the nearest integer gives

$$T_{\Delta w} = 3T$$

Then, for the processing of $A_3(nT)$, every third sample is selected. Utilizing 10% of the samples will yield a frequency resolution given by

$$f_{res} = \frac{1}{233} = 10 \text{ cps}$$

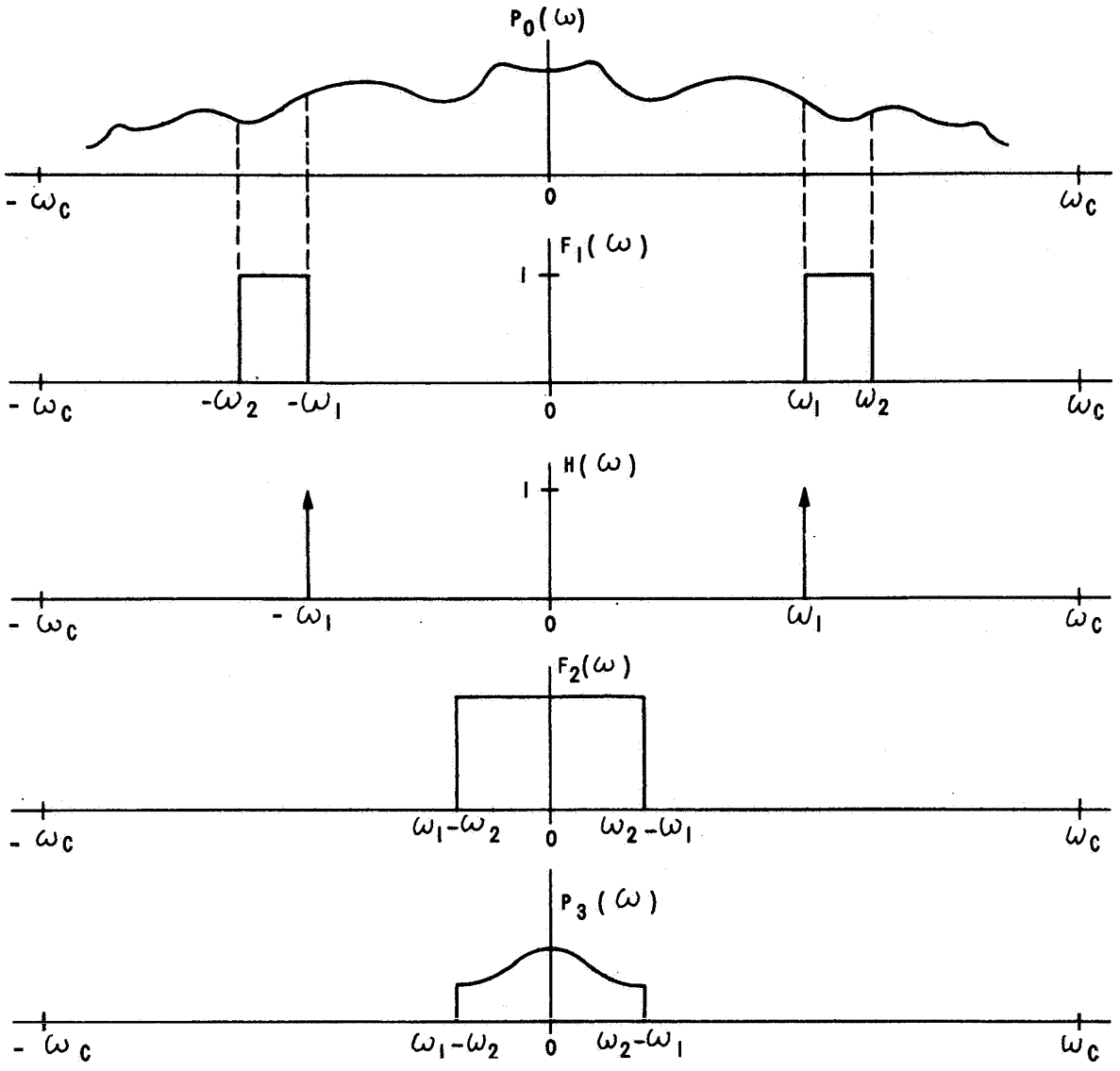


Figure 6 SAMPLE SPECTRUM

Keeping the cut-off amplitude of the digital filters in the time domain within 5% requires the use of twenty weights of $f_2(n)$ and forty weights of $f_1(n)$ per side. The times to filter, heterodyne, filter and compute the correlation and spectrum are summed to get the total computer time. Summing yields approximately 502 normalized computer operations (average **FAD** time = 1.0) for the computation of the unweighted power spectrum of $x_3(nT)$ via digital heterodyning. To compute $P_3(\omega)$ with the same resolution without heterodyning, 1921 normalized computer operations would be required. Then for this particular example, the time savings factor is 502/1921 = .26.

AUTOCORRELATION FUNCTIONS OF STATIONARY PROCESSES

With the advent of the general purpose digital computer many users have developed general purpose data reduction programs. One of the most useful results of these processors has been the computation of power spectral densities. Analog processors usually use filters to directly obtain the spectra. Digital methods usually employ the technique of first computing the correlation function of the process, then the power spectral density is obtained by (digitally) performing the Fourier transform of the correlation function.

The disadvantage of this technique is the tremendous number of multiplications and additions usually required. For this reason, new techniques have been studied which, when properly applied, will reduce the computational time. These techniques and methods of application are given in this section of the report.

The Usual Correlation Computation. - Assume that a zero-mean sequence of N uniformly-spaced samples x_n is available. Common practice is to compute a correlation function at lag m by the process:

$$R_1(m) = \frac{1}{N} \sum_{n=1}^{N-m} x_n x_{n+m} \tag{35}$$

or by the process

$$R_2(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} x_n x_{n+m} \tag{36}$$

Where x_n has the full precision of the data samples. These correlators are referred to as the "full-precision" correlators. Note that in either case for each m , $(N-m)$ data multiplications are required, together with $(N-m)$ additions. Neglecting the scaling for each m in the R_2 method, a total of

$$\sum_{m=0}^{m_{max}} (N-m) = N(m_{max}+1) - \frac{m_{max}(m_{max}+1)}{2} \tag{37}$$

multiplications and the same number of additions are required. For large N and m_{max} as many as 10^7 multiply and add times may be required. Such large numbers of operations make computational times quite long even for the very high speed general purpose computers of today. Hence, methods which employ only fast logic and additions should be quite useful.

Alternate Schemes. - A first alternate scheme is to compute

$$R_3(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} (\operatorname{sgn} Y_n)\, Y_{n+m} \tag{38}$$

where sgn is the signum* function. In this scheme, multiplications are avoided. The method has been used in a cross-correlation context [1, 2] in a plant-identification scheme. During the remainder of this report, R_3 has been termed the "half-polarity correlation."

A second alternative scheme is to compute

$$R_4(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} (\operatorname{sgn} Y_n)(\operatorname{sgn} Y_{n+m}) \tag{39}$$

In this scheme, multiplications are avoided, and fast-assess storage is reduced. Applications of this system have been suggested for spectrum analysis for ECM, and for detection devices [3, 4, 5]. This scheme has been termed the "polarity coincidence correlator."

A third alternative scheme is to compute

$$R_5(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} \left\{ |Y_n + Y_{n+m}| - |Y_n - Y_{n+m}| \right\} \tag{40}$$

This scheme avoids multiplication in a manner similar to the "quarter-square multiplier" concept used in analog multipliers. It is the basis for the transformer-rectifier-dc-output phase detector.[6] Applications to nonsinusoidal waves have been discussed.[5, 7] This scheme has been termed the "linear rectifier correlator."

*signum $(z) = +1$ when $z > 0$
$= 0$ when $z = 0$
$= -1$ when $z < 0$

Accuracy of the Alternate Schemes - Gaussian Noise. - When the sequence y_n represents samples from a stationary gaussian random process, the alternate scheme $R_3(m)$ is, for large N, an unbiased estimator of the true correlation function, within a known multiplicative constant.[8,9,10] Not only is $R_3(m)$ an unbiased estimator, but the standard deviation in the output is not much larger than that for the multiplier-square correlator in certain evaluated cases.[5]

The alternative scheme $R_4(m)$ does possess a known, arcsine-shaped bias.[10] This bias, which can be computed by standard techniques, can in principle be removed. However, it is possible that for spectrum analysis the bias may not seriously degrade the spectrum. With Markov gaussian noise, for example, the spectrum is distorted less than 2 db, as evidenced by a power series expansion of the arcsine function. Again, little statistical efficiency is lost.[4,5] However, little computational speed advantage of R_4 over the R_3 correlator, can be realized unless special purpose computers are used.

Accuracy of the Alternative Schemes: Arbitrary Inputs. - Little can be said in general for the accuracy of the alternative schemes in the arbitrary-input case. It can be observed that for a sinusoidal input, for $R_3(m)$ the correct cosine correlation function results.* Analysis performed here indicates that if the arbitrary signal has a large component of gaussian noise, the correlators R_3, R_4, R_5 can be made to possess small bias. This latter effect is similar to the effect of dither in nonlinear control systems.

Half-Polarity Correlation for a Gaussian Random Process

If the parent distribution of the stationary random process y_n is bivariate Gaussian, one can compute the expected value of the half-polarity correlation function

$$R_3(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} (\operatorname{sgn} y_n)\, y_{n+m} \tag{41}$$

The correlation function of a stationary process y_n is defined as

$$R(m) = E\left[y_n\, y_{n+m} \right] \tag{42}$$

* This is easily proved by expanding the signum function into its Fourier series; only the fundamental contributes to the average.

where $E[\ ]$ is the expected value. Now the expected-value of the estimator $R_3(m)$ is

$$E[R_3(m)] = E\left[\frac{1}{N-m}\sum_{n=1}^{N-m}(\operatorname{sgn} x_n)\, x_{n+m}\right]$$

$$= \frac{1}{N-m}\sum_{n=1}^{N-m} E\left[(\operatorname{sgn} x_n)\, x_{n+m}\right]$$

$$= E\left[(\operatorname{sgn} x_n)\, x_{n+m}\right] \quad \text{, for } x \text{ stationary} \tag{43}$$

Let $u = x_n$, $v = x_{n+m}$

then

$$E[R_3(m)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\operatorname{sgn} u)\, v\, p(u,v)\, du\, dv \tag{44}$$

where $p(u,v)$ is the joint density function of u and v . Given that the x process had zero mean and normalized correlation of

$$\rho = \rho(m) = \frac{R(m)}{R(0)}$$

$$p(u,v) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\, e^{-\frac{u^2-2\rho uv+v^2}{2\sigma^2(1-\rho^2)}} \tag{45}$$

where $\sigma^2 = R(0)$

Substituting $p(u,v)$ into the indicated double integral yields

$$E[R_3(m)] = \int_{-\infty}^{\infty} du\, \frac{\operatorname{sgn} u}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \frac{v}{\sqrt{2\pi\sigma^2(1-\rho^2)}}\, e^{-\frac{u^2-2\rho uv+v^2}{2\sigma^2(1-\rho^2)}} dv \tag{46}$$

The exponent in the second integral can be written as

$$-\frac{v^2-2\rho uv+u^2}{2\sigma^2(1-\rho^2)} = -\frac{v^2-2\rho uv+\rho^2u^2}{2\sigma^2(1-\rho^2)} + \frac{\rho^2u^2-u^2}{2\sigma^2(1-\rho^2)}$$

$$= -\frac{(v-\rho u)^2}{2\sigma^2(1-\rho^2)} - \frac{u^2}{2\sigma^2} \tag{47}$$

substitution yields

$$E[R_3(m)] = \int_{-\infty}^{\infty} du\ \operatorname{sgn} u\ \frac{e^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \frac{v\, e^{-\frac{(v-\rho u)^2}{2\sigma^2(1-\rho^2)}}}{\sqrt{2\pi\sigma^2(1-\rho^2)}}\, dv \tag{48}$$

Using the identity

$$E[x] = \int_{-\infty}^{\infty} x\, p(x)\, dx = \int_{-\infty}^{\infty} \frac{x\, e^{-\frac{(x-m)^2}{2\sigma_x^2}}}{\sqrt{2\pi\sigma_x^2}}\, dx = m \tag{49}$$

with $\sigma_x^2 = \sigma^2(1-\rho^2)$ and $m = \rho u$,

$$E[R_3(m)] = \int_{-\infty}^{\infty} \frac{(\operatorname{sgn} u)\, e^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, \rho u\, du$$

$$= \rho \int_{-\infty}^{\infty} \frac{u \operatorname{sgn} u}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{u^2}{2\sigma^2}}\, du \tag{50}$$

$$= \rho \int_{-\infty}^{\infty} |u|\, \frac{e^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, du$$

$$= 2\rho \int_{0}^{\infty} \frac{u\, e^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\, du$$

performing the integration by letting $r = \frac{u^2}{2\sigma^2}$

$$E[R_3(m)] = \frac{2\rho}{\sqrt{2\pi\sigma^2}} \int_0^\infty e^{-r} \sigma^2 dr$$

$$= \sqrt{\frac{2}{\pi}}\, \sigma \rho(m)$$

$$E[R_3(m)] = \sqrt{\frac{2}{\pi}}\, \frac{R(m)}{\sqrt{R(0)}} \tag{51}$$

Hence for a stationary gaussian process with zero mean and normalized correlation of $\rho(m)$, the $R_3(m)$ is a biased estimator of the true correlation function. The relationship between the normalized correlation function ρ and ρ_3 is computed as follows:

$$E[R_3(m)] = \sqrt{\frac{2}{\pi}}\, \frac{R(m)}{\sqrt{R(0)}} \tag{52}$$

$$E[R_3(0)] = \sqrt{\frac{2}{\pi}}\, \frac{R(0)}{\sqrt{R(0)}} \tag{53}$$

Since

$$E\left[\frac{R_3(m)}{R_3(0)}\right] \approx \frac{E[R_3(m)]}{E[R_3(0)]}$$

$$E[\rho_3(m)] \approx \rho(m) \tag{54}$$

The new normalized correlation function is therefore approximately unbiased. Since $R_3(m)$ can be computed more rapidly than the full-precision estimator, computer time can be saved if zero-mean gaussian random variables are to be processed. Note that this is also a test for a gaussian process. Other random processes also exhibit this correlation function invariance, e.g. pure sine waves or square waves with random phase. (See Ref 9).

Comparison of R_3 and R_2. - For the work report here, NASA supplied CAL with four time records which are typical of the random processes analyzed by the computation Laboratory. These records were processed by CAL using the R_2 and R_3 correlation techniques. The resulting correlation functions and power spectral densisite (PSD) for the two techniques are shown on plots in Appendix D. The PSD's were obtained by computing the Fourier transform of the corresponding normalized autocorrelation functions. (Hanning smoothing was also applied to the PSD)

Upon visual comparison, one can see that the spectra are in very close agreement. Several "numerical" comparisons were attempted during the analysis but such measures as the average, RMS, and variance of the differences between the $R_2(m)$ and $R_3(m)$ and resulting power spectra did not yield useful comparisons. Perhaps the "best" comparison is the visual examination of the plots.

Methods of Computing the Half-Polarity Correlator. - A complete digital program was written to compute in minimum time the $R_3(m)$ correlator defined as:

$$R_3(m) = \frac{1}{N-m} \sum_{n=1}^{N-m} (\operatorname{sgn} x_n)\, x_{n+m} \qquad (55)$$

A prepacking technique, written in MAP, enabled the $R_3(m)$ to be computed without any "IF" statements, etc., so that only $(N-m)$ ADD times are required per lag. To further reduce the computation time, the above summation for R_3 was replaced with the equivalent Stieltjes summation (after the Stieltjes integral). In effect the x_{n+m} are replaced with the accumulative $X_{n+m} = \sum_{j=1}^{n+m} x_j$ and $\operatorname{sgn} x_n$ is replaced by

$$\delta_n = \operatorname{sgn} x_n - \operatorname{sgn} x_{n-1} = 0, \pm 2 \qquad (56)$$

When $\delta_n = 0$ it is unnecessary to compute the "product," $\delta_n \cdot X_{n+m}$, thus an ADD time is eliminated for each $\delta_n = 0$. The flow chart of the $R_3(m)$ correlator is given in Appendix E. A complete theoretical treatment of the correlation function of quantized date (including R_3) is contained in Reference 21, and indicates that for 80 to 90 per cent of experimental data, δ_n is zero. For the NASA data, summarized in Table 1, about half the δ_n were equal to zero.

Table 1 summarized the computer time required to calculate the PSD functions of the NASA data records. Note that the computation of the $R_2(m)$ correlation function has been minimized somewhat because Fortran

programming was not used. A detailed comparison of the resulting PSD's can be seen in Appendix D, as mentioned in the above discussion.

TABLE 1. COMPARISON OF COMPUTER [a] TIME

FUNCTION	AVERAGE[b] TIME - SECONDS	
	$R_2(m)$ [c]	$R_3(m)$ [d]
CORRELATION[e]	238	11
PSD[f, g]	56	56
MISCELLANEOUS: Input, Output, Cosine Table, and Hanning	16	16
TOTAL TIME [h]	310	83

[a] IBM 7044 - 2 usec/cycle.

[b] For 4 example records furnished by NASA.

[c] Using "MAP" programming, this R_2 correlator takes approximately one-half the time of the equivalent Fortran compilation.

[d] Half-polarity correlator with prepacking and Stielje's Integration technique.

[e] For 700 lags and 7000 data points.

[f] For 700 values of frequency.

[g] The 56 sec. can easily be reduced to 14 sec. with a more "optimum" computation of the Fourier transform.

[h] Times can be reduced to 268 and 41 sec. respectively with the more "optimum" PSD computation.

Correlation of Signal Plus Noise

In the above sections, the R_3 correlator was shown to be an unbiased (or **known** bias) estimator of the true correlation function for certain random or deterministic processes; for example, a gaussian bivariate process and a single sinusoidal signal. It has also been demonstrated and summarized in the previous section of this report that the R_3 correlator can have as much as 20:1 computer time advantage over a carefully programmed R_2 (full precision) correlator. It becomes apparent that computer time can be minimized if the R_3 correlator can be used for a larger class of random processes. Furthermore, it would be necessary to determine, by a rapid testing procedure, to what class a given signal belongs. Not much effort has been applied to determine a good testing procedure however, this area of research is recommended for future study.

On this project more effort was expended io increase the class of random processes for which the R_3 estimator is unbiased. An analysis has been performed to compute the bias of the R_3 correlator for all stationary processes. The technique is to add uncorrelated noise (white) to the random process before using the R_3 correlator. One processes the signal plus noise rather than signal alone. Let the processing of signal plus noise by the R_3 correlator be designated by R_3' . Then the expected value* of this correlator is given by

$$E\left[R_3'(m)\right] = \sqrt{\frac{2}{\pi}}\,\frac{\sum_{k=1}^{N-m}\sqrt{R_n(0)}\;e^{-\frac{f_{k+m}^2}{2R_n(0)}}}{N-m}\,\delta(m) + \sqrt{\frac{2}{\pi}}\,\frac{1}{N-m}\sum_{k=1}^{N-m} f_k \int_0^{\frac{f_{k+m}}{\sqrt{R_n(0)}}} e^{-\frac{x^2}{2}}\,dx \qquad (58)$$

where

$$\delta(m) = \begin{cases} \quad ; & m = 0 \\ 0 \quad ; & m \neq 0 \end{cases}$$

f_n = samples of signal

$\sqrt{R_n(0)}$ = rms value of white noise added

The above relation indicates that the $R_3'(m)$ correlator yields an unbiased estimate of the $R_2(m)$ correlator when the noise strength is sufficiently

*See Appendix F for the derivation of this expected value.

large compared to the signal sample f_m and $m \neq 0$. For zero lag, $m=0$, an additional term appears and is a direct consequence of the addition of white noise. This term will produce a bias in the PSD function; however, the bias is of known magnitude and can be compensated for.

With the knowledge that the expected value of the R_3' correlator is that of the R_2 correlator for any process f_m under the conditions cited above, the variance of the error was then investigated. An approximate expression for the normalized variance of the error for non-zero lag is given* as

$$E\left[\left(R_2(m) - \sqrt{\frac{\pi}{2} R_n(0)}\; R_3'(m)\right)^2\right] \Big/ R_2^2(0) \simeq$$

$$\left[\frac{R_2(m)}{R_2(0)} - \frac{\sqrt{R_n(0)}}{R_2(0)} \frac{1}{N-m} \sum_{k=1}^{N-m} f_k \int_0^{\frac{f_{k+m}}{\sqrt{R_n(0)}}} e^{-\frac{\chi^2}{2}} d\chi\right]^2$$

$$+ \frac{\pi}{2} \frac{1}{N-m} \left[\frac{R_n(0)}{R_2(0)}\right]^2 \qquad (59)$$

The first term of equation (59) may be looked upon as a bias term and the second term as the noise-in-output contribution. Increasing the noise strength will reduce the bias tkrm; however, it will increase the noise-in-output contribution.

Test of R_3 on Signal Plus Noise. - In order to show the effect of adding noise to the signal before processing with the R_3 correlator, a test case was generated. The test signal was a periodic function

$$y(t) = \sin 2\pi(700)t + \tfrac{1}{2} \sin 2\pi(2100)t \qquad (60)$$

which has zero mean. Since x is a distorted sinusoidal function, the R_3 correlator, used directly to estimate the correlation function of x , will not provide a "good" estimate.

As described above, one should add uncorrelated noise to the signal x before processing with the R_3 correlator. For the sample signal selected the variance of noise added was set equal to the variance of x . (This variance of x has to be precomputed). To

*The assumptions and derivation of this expression is not given in this report.

demonstrate the results, the correlation functions and the respective power spectra were computed as follows:

(1) The full precision (R_1) correlation function of $x(t)$. This should give the "best" estimate of the true spectrum.

(2) The half-polarity (R_3) correlation function of $x(t)$. This should give a "poor" estimate of the spectrum.

(3) The half-polarity (R_3) correlation function of $x(t)$ plus the uncorrelated noise. The resulting spectrum should be "better" than (2), but not as "good" as (1).

The resulting spectra are shown in Figure 7. Although no test for "goodness" is given, a visual examination of the plots certainly demonstrates the useful effect of adding noise before using the half-polarity correlator. Note how smooth the spectrum of $P_3(n+x)$ is compared to $P_3(x)$. Also note the accurate estimate of the power at the two frequency components of $x(t)$, namely 700 and 2100 cps.

High Speed Techniques of Computing the Full-Precision Correlator

Rather than sacrifice the full precision of the sampled sequence x_n for computer speed, various methods of calculating the $R_1(m)$ or $R_2(m)$ have been investigated. One simple method is to use "tighter" programming, that is programming in basic machine "language" to avoid the inefficiencies of a general purpose compiler such as FORTRAN. Further, "optimum" use of the commands available with various machine types can help to reduce the computational time. For example, the variable word. multiply and accumulate (VMA) command of the IBM 7044 can be used to great advantage.

Other schemes employ special (digital) computing machinery to calculate the correlation functions. One method is to use residue numbers [31, 32, 33, 34]. This scheme was studied for this contract. The residue number system is of particular interest because the arithmetic operations of addition, subtraction, and multiplication may be executed in the same period of time without the need for carry. The equipment proposed in the literature suggests very rapid computation of correlation functions. Unfortunately, the scheme could not be conveniently programmed or a general purpose digital computer, so extensive study, for comparison to other methods, was not completed during this contract. It is suggested that where one desires to calculate the correlation funtions in almost real time, this special purpose technique should be considered.

Other programming techniques use the table look-up to great advantage. One such method was recently proposed in a letter to the

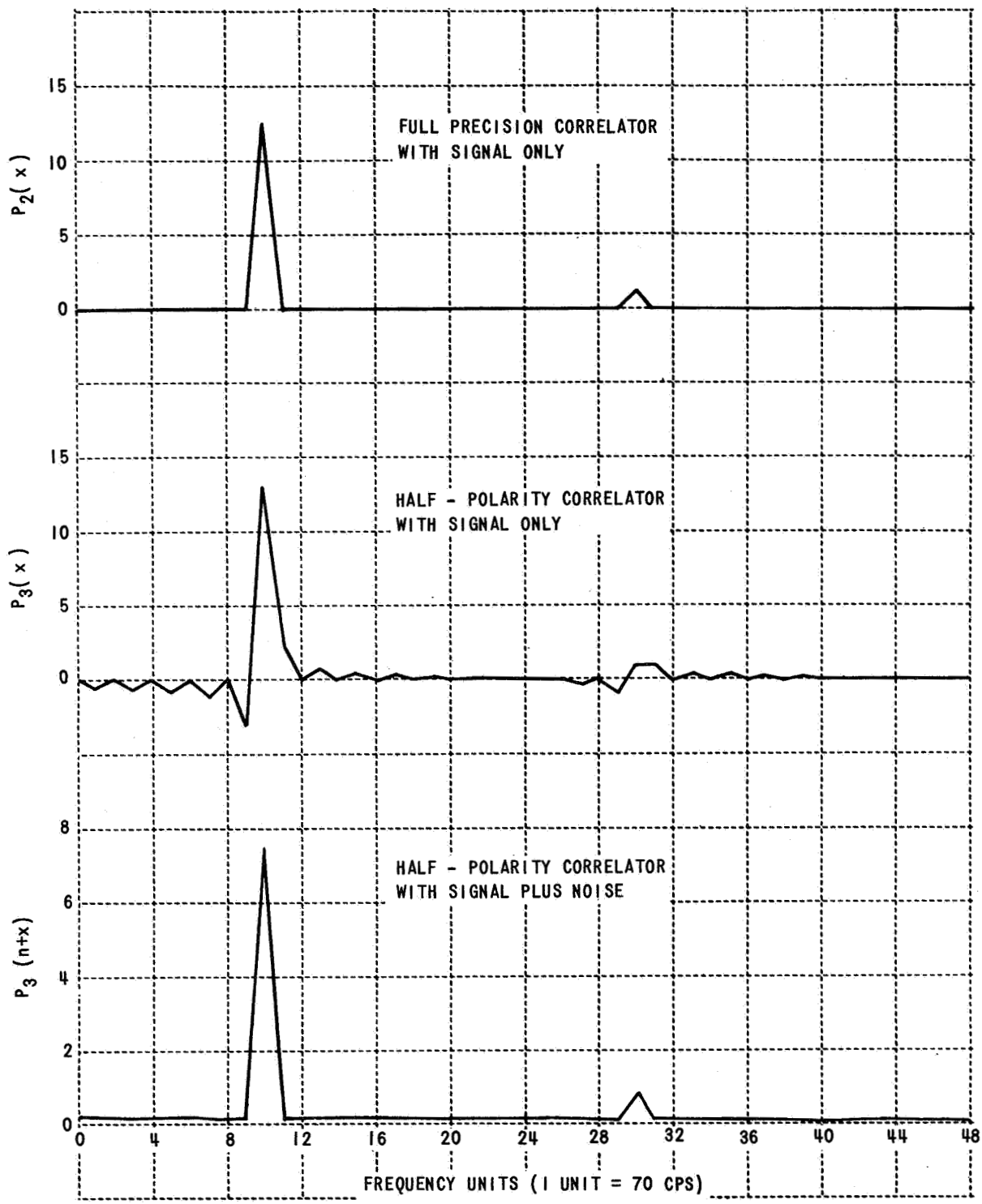


Figure 7 TEST OF R_3 CORRELATOR WITH SIGNAL PLUS NOISE

editor of the Communication of the ,Association for Computing Machinery (see reference 7). This method suggests the use of table look-up coupled with the well-known "quarter square" method used for analog multipliers.

Further techniques suggest the use of pre-sorting and packing of the data, which usually eliminate many multiplications because of the "sameness" of the quantized samples. One such method has been designed by W. D. Fryer of CAL. This method is summarized in the following sections of this report.

The Pre-sort Correlation Method (PSCM)

In the implementation of the pre-sort correlation method (PSCM) described in this section, numerical data are quantized, sorted, then transformed into a set of executable computer instructions for rapid calculation of the correlation sums

$$C_k = \sum_{i=1}^{N-k} A_i B_{i+k} \quad ; \; k = 0, 1, 2 \cdots L \qquad (61)$$

Here $A = \{ A_i , i = 1, 2, \cdots N \}$ **dnd** $B = \{ B_i , i = 1, 2 \cdots N \}$ are two* data sequences of length N; k is the so-called lag number; L is the maximum lag number; and the C's may be thought of as unnormalized correlation function values.

In order to use the method, data in one of the sequences (A , say) are quantized, so that magnitudes of these data assume at most N_v different values. The method provides a speed advantage when $N_v << N$. The ratio $\rho = N/N_v$ is a measure of speed improvement factor over conventionally coded methods, in a sense to be defined exactly later.

Conventional Coding: Timing. - In order to set a reference for judging the savings in number of instructions and in speed, a typical coding of the C_k sums, Eq. 61 is shown here in the language of IBM 7000-series computers:

---

*The program always treats the **A** and B sequences as logically distinct, even in autocorrelation sums where -- originally -- they may be numerically identical.

(Initialize index register 1 to N-k)

(Initialize SUM to zero)

```
LOOP    LDQ   A+N-K,1
        MPY   B+N,1    (or FMP)
        ADD   SUM      (or FAD)
        STO   SUM
        TIX   LOOP,1,1
```

(computing loop ended)

An individual term in any one of the C_k sums corresponds to a single passage through this sequence of instructions. In terms of IBM 7044 timing values, the amount of computing time required would average about 20 cycles (40 microseconds) for the fixed-point instructions ,'just slightly more for the floating-point instructions. Thus, evaluation of say C_0 requires about 20N cycles.

PSCM: timing. - Before describing the PSCM itself, the manner in which its speed advantage arises is discussed. An add instruction must be executed the same number of time (N times for C_0) as in conventional coding. However, it is approximately true that all other instructions are executed not N times, but about $N_v = N/\rho$ times (where $\rho = N/N_v$ is the "improvement factor" previously mentioned). The value $\rho = 7$ will arise later($N = 7000$, $N_v = 1000$);using this value for illustration, the evaluation of C_0 , using the PSCM, would require roughly

$$2N + 18\,N/\rho \approx 4.57\,N \quad \text{cycles} \qquad (62)$$

compared with $20N$ cycles for the conventional coding method of the previous paragraph.

Description of the PSCM: Example. - Logic of the PSCM is demonstrated here by means of a simplified, illustrative problem.

Suppose the A data, after quantization if necessary, assume only values $\pm 1, \pm 2$, or ± 3 according to this partial table:

$i =$	1	2	3	4	5	6	7	...	N-1	N
$A_i =$	3	2	-3	-3	1	-2	-1	...	-1	2

Conventional coding for the zero-lag sum, C_0 , is the direct implementation of the sum

$$C_0 = 3B_1 + 2B_2 - 3B_3 - 3B_4 + 1B_5 - 2B_6 - 1B_7 + \cdots - 1B_{N-1} + 2B_N , \qquad (63)$$

with, of course, N multiplications, (algebraic) additions, and auxiliary instructions.

Coding the pre-sort method, in contrast, directly implements this numerically equivalent formulation:

$$C_5 = 3(B_1 - B_3 - B_4 + \cdots) + 2(B_2 - B_6 + \cdots B_N) + 1(B_5 - B_7 + \cdots - B_{N-1}) \quad (64)$$

The dominant feature, of course, is that an algebraic sum of $N/N_v = N/3$ terms (on the average) may be accumulated before a multiplication is necessary. Or, put differently, only three multiplications are necessary, rather than N

It is not obvious, from this example, that the auxiliary instructions (the LDQ, the STO and the TIX instructions) of the conventional coding are also largely removed. The PSCM can be coded, however, to accomplish this additional time savings.

Programming Logic for the PSCM. - Although the previous example demonstrates the principle by which a time savings may be achieved, the actual programming requires careful attention to details.

First, data in the A sequence may require quantization.

Second, these data must then be sorted into blocks, each block corresponding to a distinct value of magnitude. Data inserted into these blocks must carry information about their locations in the original sequence.

Third, data within a block must be converted from numerical to instruction words, either ADD (if data positive) or SUBTRACT (if data negative). These instructions must also be supplied with proper address values (determined from locations in original sequence), and with an index register tag. Each of these blocks, at execution time, will form an algebraic subtotal such as represented in equation (64) by the parenthetical expressions.

Fourth, instructions must be inserted between blocks to multiply (block "value" times subtotal) and to update the over-all sum.

Fifth, "header" and "trailer" instructions must be provided for initializing an index register (in accordance with lag number and absolute location of the B sequence), for initializing sums, for storing results, and for terminating calculation at the proper time.

Figure 8 shows what typical sections of storage would contain, after the conversion of the A data. Assume for definiteness. that A data is used in the previous example with $N = 1000$. Let these

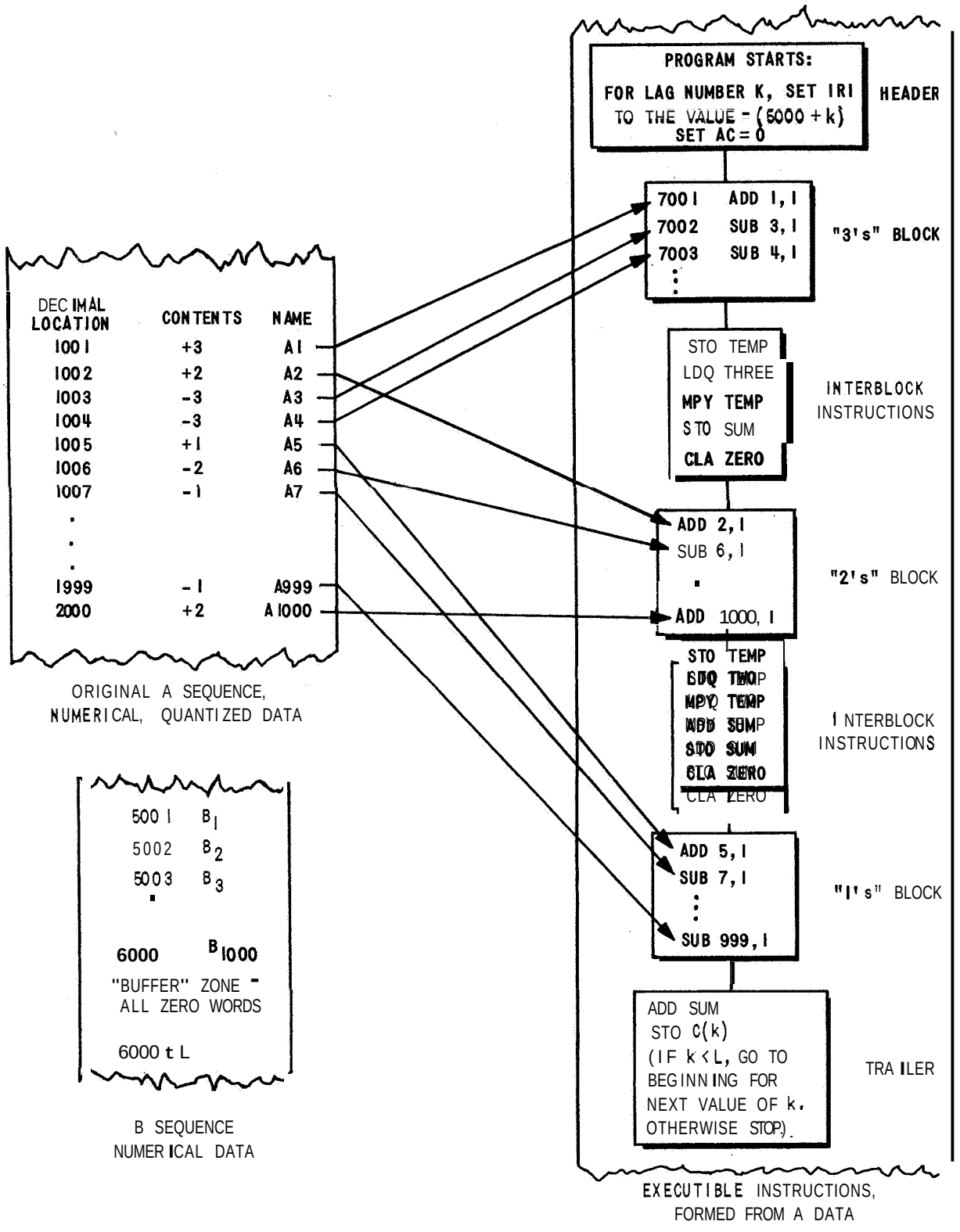


Figure 8 REPRESENTATIVE DIAGRAM OF STORAGE

1000 A-data be stored in locations 1001 through 2000 (decimal); the 1000 B-data be stored in locations 5001 through 6000; and assume something more than 1000 words, beginning near 7001, reserved for formation of the modified A-data -- the executable instructions of the program.

Suppose now that the correlation sum for lag number $k = 0$ is desired. The program at the right-hand side of Figure 8 would process:

Index register 1 set to the value -5000.

Accumulator zeroed.

The next instruction reads ADD 1, 1.

Its effective meaning is: ADD 5001,

that is: ADD B_1

Similarly, the next few instructions are effectively:

SUB B_3

SUB B_4

Thus, the "3's" block forms

$$B_1 - B_3 - B_4 + \cdots ,$$

The first parenthetical expression of equation (64). The interblock sequence forms 3 $(B_1 - B_3 - B_4 + \cdots)$, and stores this partial sum away in SUM. Analogously, the "2's" block forms the subtotal

$$B_2 - B_6 + \cdots + B_{1000} ,$$

the second parenthetical expression of equation (64). The interblock instructions multiply this subtotal by 2 and increments SUM accordingly.

Finally, the "1's" block forms the subtotal

$$B_5 - B_7 + \cdots - B_{999} ,$$

and final incrementing of SUM to give C_0 is executed in the "trailer".

The program would now return to the start, with k set equal to 1, and index register 1 initialized to the value -5001. The following sequence of instructions will then form

$$3(B_2 - B_4 - B_5 + \cdots) + 2(B_3 - B_7 + \cdots + B_{1001}) + (B_6 - B_8 + \cdots - B_{1000}),$$

which is the correct value for the C_1 correlation sum, except that beyond the nominal extent of the original B array -- has appeared. Thus the sum will be correct if the location B_{1001}, contains zero. Generally, this over-extension of the original B array will occur, eventually, for L locations, and may most easily be handled (with only small loss of time, since the number of lags, L, is usually a small percentage of the total data record length, N) by adding zero-valued buffer words at the end of the B array.

OPTIMUM SPECTRAL SMOOTHING FOR STATIONARY STOCHASTIC PROCESSES

The criteria of bestness which is used to obtain optimum smoothing weights is given by the following performance index:

$$J = \int_{-\omega_c}^{\omega_c} \left| \Phi_a(\omega) - \Phi_r(\omega) \right|^2 d\omega \qquad (65)$$

In equation (65) the power spectral density function $\Phi_a(\omega)$ is the spectrum which is obtained when a given correlator and lag window pair are selected to operate upon the data to be analyzed. Taking the inverse sequence transform of $\Phi_a(\omega)$ yields

$$\mathcal{Z}^{-1}\left\{\Phi_a(\omega)\right\} = D_j R_j \qquad (66)$$

where in (66) D represents the lag window and R the correlation function.

The true underlying power spectra of the process being analyzed is specified in equation (65) by $\Phi_r(\omega)$ and in general is unknown. Smoothing weights are found by minimizing the expected value of J with respect to D_j. Results of this analysis, details of which are contained in Appendix G, is given by the following expression for the smoothing weights.

$$D_j = \frac{\phi_j E[R_j]}{E[R_j^2]} \qquad j < N \qquad (67)$$

where

ϕ = true correlation function

N = number of data points

Restrictions on the index in equation (67) are implied by limiting the class of correlators being considered to those which yield a zero value of correlation when the lag value equals the number of data points. In particular, if the R_2 correlator is selected, the expression for the optimum

weights in equation (67) becomes

$$D_J = \frac{\phi_J^2}{E[R_J^2]} = \frac{\phi_J^2}{\phi_J^2 + VAR[R_J]} \qquad J < N \tag{68}$$

If the ratio of the correlator variance to the true value of correlation squared is defined to be the normalized variance ν, equation (68) can be used to compute the assumed normalized variance ν as a function of the lag window selected, It would be of interest to obtain plots of this normalized variance for some of the more "standard" lag windows in use. In practice, the normalized variance is often assumed to be infinite after a lag value of about ten percent of the record length.

A plot of the assumed normalized variance should in fact accompany each power spectral density function since the latter is computed with the aid of the former. The great number of questions which would undoubtedly arise, from those interested solely in a final spectral density function would most likely preclude the effective use of such a plot'.

With the assumption that the process under investigation is Gaussian equation (68), as shown in Appendix H, may be expressed as

$$D_J = \frac{\phi_J^2}{\phi_J + \frac{1}{(N-J)^2} \sum_{k=-(N-J-1)}^{N-2-1} (N-J-|k|)(\phi_k^2 + \phi_{k+J}\,\phi_{k-J})} \qquad J<N \tag{69}$$

The relation given by (69) illustrates the need for appropriate prewhitening of the spectrum when a "standard" lag window, which becomes zero well before the maximum possible lag, is utilized. The effect of prewhitening is to reduce the correlation $\phi_J\ (J \neq 0)$ at a given lag and thus dictates the use of a lag window D which decreases with increasing lag J. This is so since the numerator term in (69) is ϕ_J^2 while the denominator contains $\phi_0^2/(N-J)$. When increased accuracy of the spectral estimates is the main concern however, it appears more advantageous to generate the smoothing weights in the manner later described rather than select a standard window.

The arbitrary selection of a given set of smoothing weights, when adequate prewhitening is omitted, is to a degree forcing the shape of the final spectrum to take on certain aspects. It would be informative to obtain a solution (or envelope of solutions) to equation (69) when the weights D_J are selected to be Hanning, Hamming, Bartlett, or any one of the other standard forms.

Adequate prewhitening implies some knowledge of the spectrum under consideration. This knowledge may well represent the summation of previous experience with a given process or perhaps information obtained from a pilot estimation. However obtained, this best knowledge can be used in conjunction with equation (69) to produce a set of optimum smoothing weights. Generation of the weights in this fashion will remove the necessity of prewhitening operations but will in general increase the maximum lag value of the correlation taken into consideration.

DETERMINISTIC DATA PROCESSING

Consider the generation of a power spectral density function, for a single record $x(t)$ of length T seconds, by the following formula (see Reference 23, p 50).

$$\phi(\tau) = \frac{x(t) * x(t)}{T}$$

$$\Phi(\omega) = \mathcal{F}\{\phi(\tau)\} \qquad (70)$$

In equation (70), it is assumed that the signal is zero outside of the observation interval of T seconds regardless of whether the signal actually exists outside this interval or not. Application of equation (70) will be referred to as deterministic data processing. This classification follows logically since in (70) there is no notion of an ensemble of signals, nor signals of infinite durations. The average power between two frequencies ω_1 and ω_2 for the observation period is also precisely given by:

$$\text{Avg. power within frequency band } \omega_2 - \omega_1 = \frac{1}{2\pi}\int_{\omega_1}^{\omega_2} \Phi(\omega)\, d\omega$$

Reasons for the selection of a deterministic approach to spectral analysis in any given situation vary. As a hypothetical case, consider the testing of a gear box which is known to be defective. The analyst who receives the piece of machinery is interested in determining the faulty gear before discarding the gear box. Expediency dictates a series of tests with microphones while the gears are running rather than disassembly and inspection. The record length is selected in accordance with the frequencies of interest. Since the concept of infinite record length cannot possibly yield additional information, the signal is processed deterministically.

The R_1 correlator discussed earlier satisfies the definition in (70). It is significant to observe that no negative power will result in the final spectrum with this approach, If in the interest of speed of computation, the correlation function is not evaluated up to the maximum lag, a unity lag window should be utilized.

NONSTATIONARY SPECTRUM ANALYSIS*

The successful characterization of a nonstationary power spectrum depends largely upon assumptions about the actual time variations of the spectrum. Restriction of the general class of nonstationary processes implies some a priori knowledge of the process or its spectrum. A priori knowledge of the time variations may enable the spectrum analyzer to be designed in an optimum manner for a particular type of nonstationary process. Of considerable interest is the vibration record of a rocket during the lift-off phase and its subsequent passage through the atmosphere which may represent a nonstationary process whose power spectrum is slowly time-varying. Analysis of this record as though it were representative of a stationary ergodic process may result in serious misconceptions as stationary processing cannot yield information about specific frequencies at particular times.

The conventional method of obtaining the power spectrum of a nonstationary process is based upon the assumption that the statistics of the signal do not vary appreciably over some interval. The greatest error in this method is that the selected intervals may be too short to yield accurate spectral estimates. If longer intervals are used, the effect is to mask possible time variations in the measured spectrum.

The principal results to be summarized in this report is the experimental verification of the usefulness of a new technique[24] for obtaining the time-varying power spectrum of a nonstationary process. The nonstationary signal that was selected for study is representative of the modulation used in "chirp" radars. The signal can best be described as a sine wave whose instantaneous frequency is slowly increasing at a constant rate. The main reason for selecting this signal is that it is representative of a wide class of nonstationary signals with the further advantage that the nature of the time variation in its power spectrum can be readily visualized.

Theoretical Considerations

The major premise of the nonstationary analysis method used in this study is that of an ensemble-average power spectrum. As a nonstationary process, consider a group of signals whose statistical description of the group is defined across the ensemble for each instant of time. If it be allowed that this statistical description across the ensemble can change with time, it is clear that a properly defined ensemble-average power spectrum will exhibit time-variations and will have the same time dimensions as the statistics of the ensemble.

$$S(t,\omega) \equiv \frac{\partial}{\partial t}\left\{\lim_{K\to\infty}\frac{1}{K}\sum_{n=1}^{K} E_n(t,\omega)\right\} \qquad (71)$$

* By Gilbert A. Gagne

where $E_n(t,\omega)$ represents the energy at time t and frequency ω of the n^{th} signal of the ensemble $y_n(t)$. The corresponding definition of the ensemble-average autocorrelation function is

$$\psi(t,\tau) \equiv \lim_{K\to\infty} \frac{1}{K} \sum_{n=1}^{K} y_n(t)\, y_n(t-\tau) \tag{72}$$

In order for the expression for the nonstationary power spectrum to apply also in the stationary case, it Is necessary to define the autocorrelation function $\psi(t,\tau)$ as an even function of the variable τ, thus

$$\psi(t,-\tau) = \psi(t,\tau) \tag{73}$$

The time-varying power spectrum can now be obtained from the time-varying autocorrelation by Fourier transformation with respect to the lag variable τ:

$$S(t,\omega) = \int_{-\infty}^{\infty} \psi(t,\tau)\, e^{-j\omega\tau}\, d\tau \tag{74}$$

Application of inverse transformation to $S(t,\omega)$ yields the time-varying autocorrelation function, the other member of the Fourier transform pair:

$$\psi(t,\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(t,\omega)\, e^{j\tau\omega}\, d\omega \tag{75}$$

The nonstationary process selected for study can be described by the following ensemble:

$$\langle y_n(t) \rangle = \langle A \cos(\omega_0 t^2 + \theta_n) \rangle \tag{76}$$

where θ_n is a random phase angle associated with each member of the ensemble. Analysis of this ensemble was performed by considering a single member which can be represented by

$$y_n(t) = A \cos[\theta(t)] + n(t) \tag{77}$$

where $n(t)$ is wideband gaussian noise whose power spectrum can be considered constant. The instantaneous frequency of $y(t)$ is given by

$$\omega_i = \frac{d}{dt}\left[\Theta(t)\right] = 2\omega_0 t \qquad (78)$$

which clearly shows the time dependency of the frequency.

The Detection Process. - In most physical situations, there is usually only one record available from which the power spectrum of the process may be derived. Since our definitions of the nonstationary autocorrelation function and power spectrum were derived from notions of ensemble-averages, some smoothing must be performed on the single record to estimate $S(t,\omega)$. The best estimate of $S(t,\omega)$ is that which is obtained by applying maximum smoothing without destroying either the desired time or frequency information. The analyzer configuration shown in Figure 9 is well suited for nonstationary spectrum analysis because it allows frequency smoothing to be separated from the time smoothing. Frequency-axis smoothing of the $S(t,\omega)$ function occurs in the predetection band-pass filters, while the time-axis smoothing occurs in the post-detection low-pass filters.

The amount of frequency-axis smoothing performed on the signal is dictated by the bandwidth of the predetection filters. Nonstationary processes of interest are characterized by continuous changes of frequency. Since any particular frequency may be present only momentarily, it is not possible to determine that frequency precisely due to the uncertainty principle relating time and frequency. Consequently, it becomes necessary to accept some smoothing along both the time and frequency axis and determine the presence of a band of frequencies rather than a single frequency.

Attempts to simulate the analyzer of Figure 9 on an analog computer revealed several practical difficulties. The most serious being that of maintaining identical bandwidths for each leg of the analyzer. Adoption of heterodyning techniques eliminated the bandwidth problem while retaining the desirable features of Figure 9. The principal feature of the analyzer shown in Figure 10 is its ability to display any line of the spectrum selected by the local oscillator.

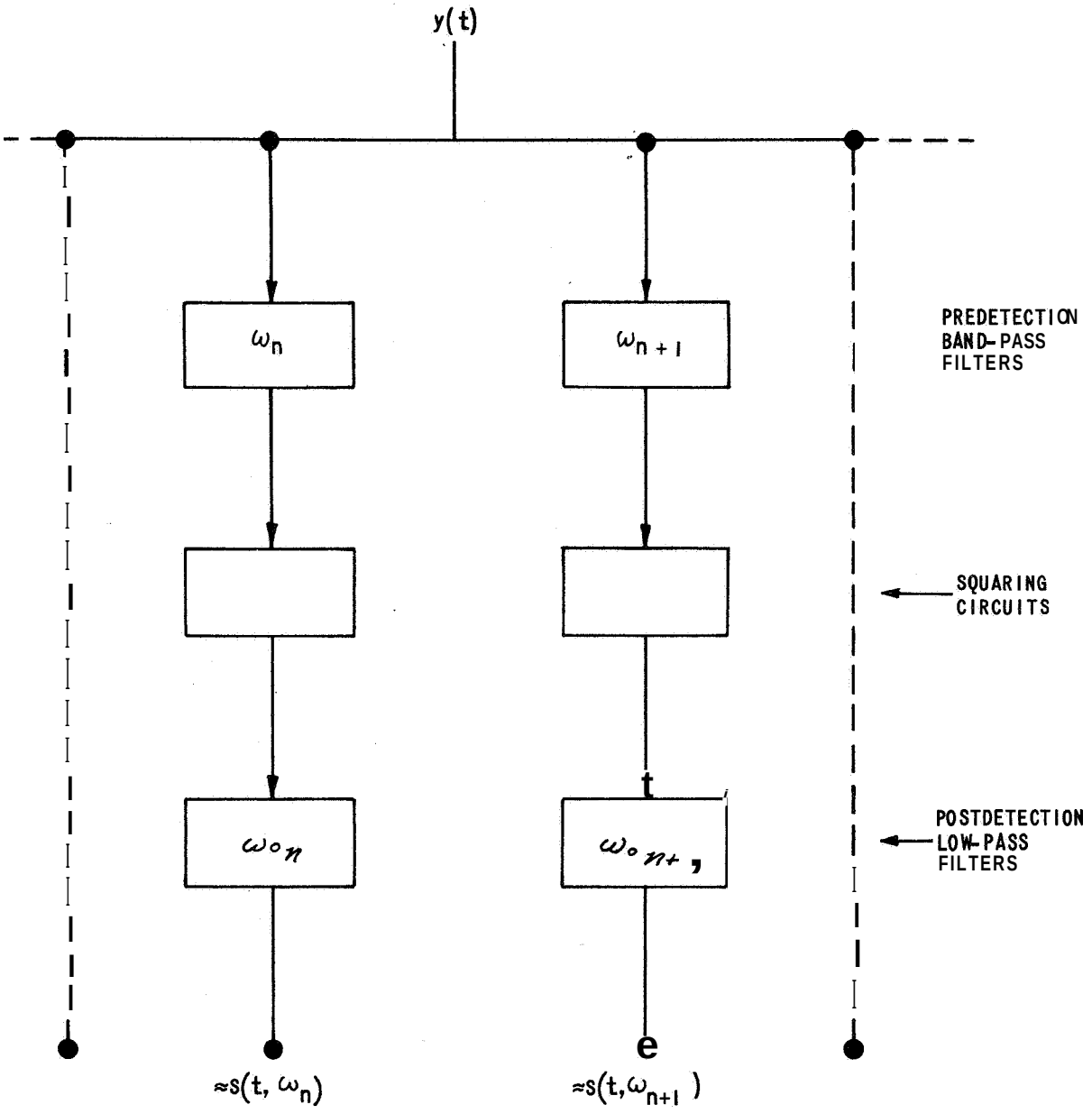


Figure 9 ELEMENTARY ANALYZER CONFIGURATION FOR NONSTATIONARY SPECTRA

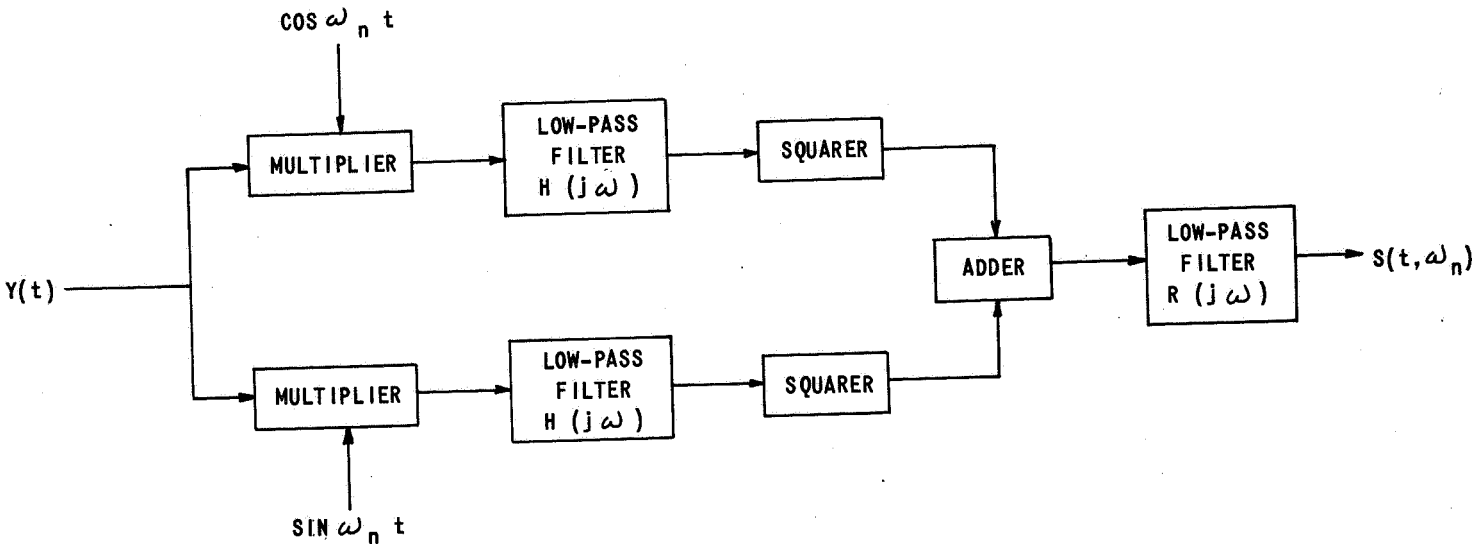


Figure 10 HETRODYNING ANA YZ≤R FOR NONSTATIONARY SPECTRA

The predetection filters shown in Figure 10 are Butterworth filters with low pass characteristics and whose transfer functions are

$$H(s) = \frac{1}{1 + .065\,s + (s/20)^2} \tag{79}$$

The damping ratio of the predetection filters was set at 0.71 and the bandwidth was 20 radians. A simple first order low pass filter with bandwidth 0.2 radians was selected for the postdetection filter.

The nonstationary test signal represented by equation (77) which was applied to the spectrum analyzer is described by

$$y(t) = e^{-\alpha t} \cos .5t^2 + n(t) \tag{80}$$

where $e^{-\alpha t}$ determines the rate of decay of the amplitude of the test signal and $n(t)$ is a wide band gaussian noise signal. Equation (80) was generated directly on the analog computer with wide band gaussian noise added from an external source. To prevent instability in the generation of the test signal, it was necessary to switch in at several points in each data run sufficient damping to reduce the signal levels in the generator thereby preventing saturation of the amplifiers. To ensure consistency this switching was performed at precisely the same point in each run. A secondary effect which resulted from the decreasing signal level was a reduction in the signal to noise ratio, the level of the injected noice being held constant during the run. The maximum value of the test signal, without noise, and which occurred at the beginning of the run had an RMS value of 7 volts. The wideband noise added to this had an RMS effective value at the output of the detector of 2 volts. The decreasing signal to noise ratio at the higher frequencies effectively demonstrated the value of optimum filtering in characterizing the time-varying spectrum of the signal.

The output of the nonstationary spectrum analyzer is shown in Figure 11. Each line (running upward to the right) represents a single run with the analyzer tuned to the particular frequency indicated at the left of that line. The same test signal (80), identically generated with wideband noise added, was applied to the input of the analyzer for each run.

The first spectral line $\omega = 0$ clearly shows the effect of the time smoothing which occurs in the post-detection filter. This line, displays the D. C, level of the signal vs. time and shows that the D. C. level eventually decreased to a small value (the noise level). A theoretically perfect nonstationary spectrum analyzer, which is unrealizable, would have displayed zero everywhere except at $t = 0$ at which point there would be an impulse equal in strength to the value of the signal at $t = 0$

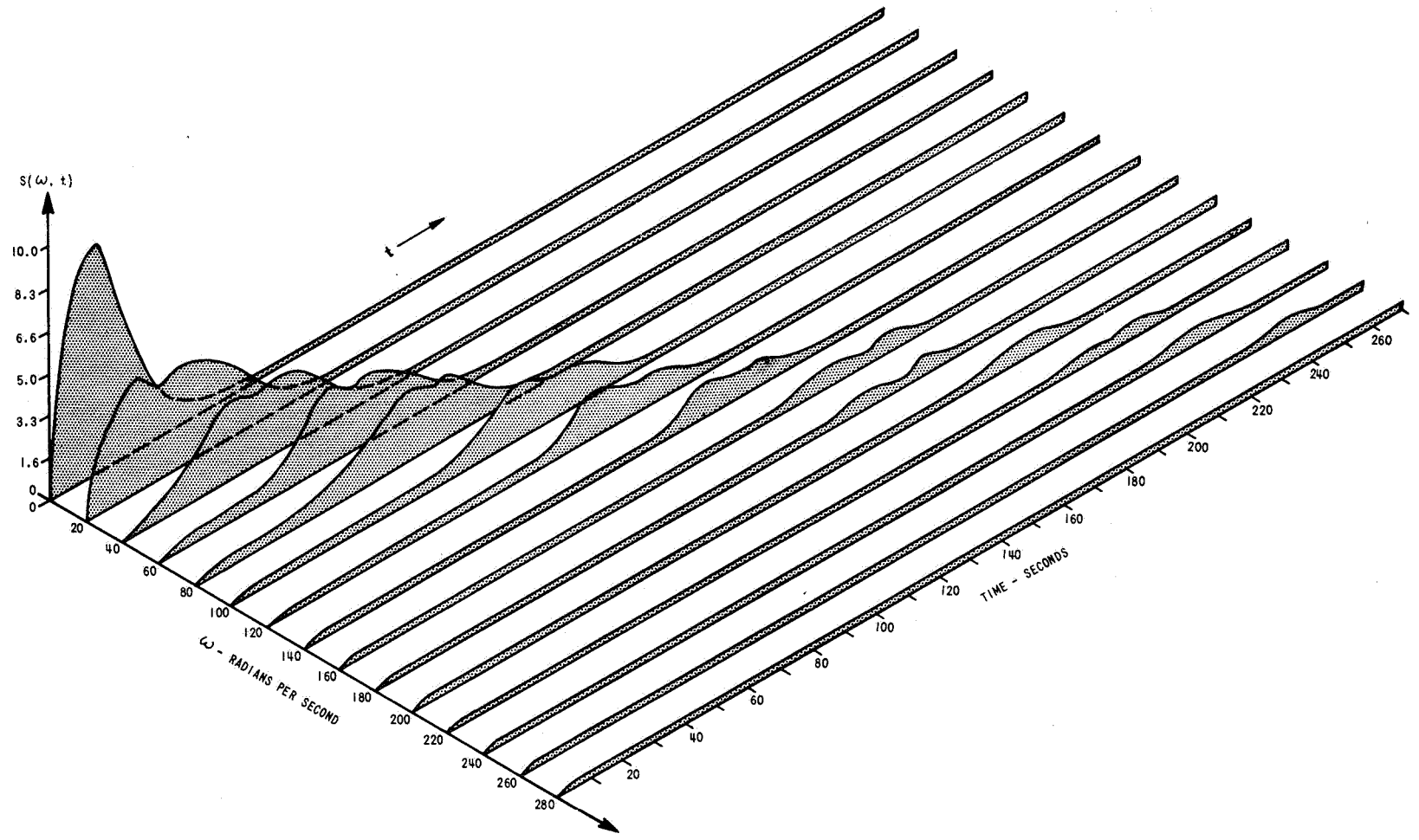


Figure 11 ANALYZER OUTPUT

The remaining spectral lines exhibit peaking in response at the proper time in each run, i.e., the 20 radian line peaks at 20 seconds, etc., with the amplitudes decreasing exponentially according to equation (80) The overall effect of Figure 11 is to show that the frequency content of the test signal is slowly shifting with time. It should be noted that a conventional stationary spectrum analyzer would only indicate that particular frequencies were present in the signal and would give no indication of when the particular frequencies appeared. It is only when an optimum amount of time and frequency smoothing is applied to a nonstationary signal that particular frequencies can be resolved in time.

Conclusions to the Spectrum Analysis of Nonstationary Processes

A new and essentially different approach to the spectrum analysis of nonstationary signals has been described. This technique utilizes both time and frequency smoothing to determine the presence of bands of frequencies. A priori knowledge of the time variations may enable the spectrum analyzer to be designed in an optimum manner for a particular type of nonstationary process. A test signal representing a wide class on nonstationary signals has been analyzed by this new technique and the results show that the method is indeed useful for determining the frequency content as well as the nature of the time variations for this class of nonstationary signals.

A THEORY OF NONSTATIONARY CORRELATION ANALYSIS*

It is generally recognized that the design of a correlator or spectrum analyzer for the analysis of nonstationary signals is a difficult task. This difficulty is the result of several problems. First, there is the problem of defining a meaningful correlation or spectral function. This function must possess characteristics which are favorable to mathematical analysis and which yield to intuitive insight. Then, there is the problem of synthesis of the analyzer which is to detect the defined nonstationary function. The synthesis procedure cannot be undertaken until after the fundamental uncertainties which shroud the nonstationary detection problem are understood. These uncertainties place definite bounds upon the accuracy with which nonstationary analyses may be performed. A third serious problem involves the fact that synthesis procedures will generally require a priori knowledge of the correlation or spectral function which is to be detected -- a most unsatisfactory situation. All of these problems must be carefully considered and overcome if a meaningful theory of nonstationary analysis is to be obtained.

Two previous approaches have been used for nonstationary analysis. They are the "ensemble approximation" approach[24] and the "short-time" approach.[25, 26] In the former, the basic concept upon which the analyzer is designed is that of approximation of an ensemble spectrum when only a single pair of time-waveforms is given. This approach places in clear perspective the uncertainties involved in the analysis of a nonstationary process, but the approach is extremely difficult to make mathematically rigorous. The latter approach makes use of an exactly detectable, deterministic spectrum or correlation function; but the detected information is of limited value, since it cannot be related to network or optimization theory. The latter approach is helpful primarily in gaining intuitive insight regarding the nonstationary process.

In this report, the underlying philosophy is closely related to that described in Reference 24. The concept of approximating the ensemble correlation function by processing a single pair of time waveforms is used. However, in this report an exact correlation analyzer synthesis procedure is developed using the methods of applied mathematics.

Definitions and Basic Relationships

The optimal synthesis procedure is initiated with the definition of the cross-correlation function and the choice of a correlation analyzer configuration which possesses special properties. The derivations which follow involve cross-correlation function analysis. Autocorrelation function analysis is then handled as a special case.

---

* This section was written by Dr. Walter W. Wierwille. The theory involved is attributed to him. An earlier version of this theory, which did not include experimental verification, can be found in Reference 35.

Let $c_1(t)$ and $c_2(t)$ be defined as a given pair of signals which are to be cross-correlated.* The independent variable t designates present time. This pair of signals is assumed to be one member pair of an ensemble of signal pairs ${}_\kappa c_1(t)$ and ${}_\kappa c_2(t)$, where κ designates the particular pair under consideration. (For convenience, the subscript designating the ensemble number of the given signal pair has been omitted.) Then the ensemble cross-correlation function is designated by the operation

$$\phi_{12}(t,\tau) \equiv \langle c_1(t)\, c_2(t-\tau) \rangle \quad ; \tau \geq 0 \qquad (81)$$

where the angular brackets represent an as yet unspecified averaging operation on the lagged products, ${}_\kappa c_1(t)\, {}_\kappa c_2(t-\tau)$.** The function $\phi_{12}(t,\tau)$ is left undefined for $\tau < 0$, since negative delays or pure predictions are generally unrealizable in a real-time solution.

The configuration which is chosen for the correlation analyzer is that shown in Figure (12). It is seen that the signal, $c_2(t)$, is first passed through a filter whose impulse response is $\ell(t)$. The output of the filter is delayed by τ seconds. Then $c_1(t)$ and the output of the delay device are multiplied together and smoothed by the linear filter whose impulse response is $h(t)$. This configuration is more general than those previously used for correlation in that the filtering operation represented by $\ell(t)$ is incorporated. Also, $\ell(t)$ and $h(t)$ are general linear filtering operations which are to be specified by the optimization process.

A remarkable property of this analyzer configuration is that it makes the t axis filtering problem and the τ axis filtering problem independent of each other. To illustrate this independence property, the output signal of the correlation analyzer is written in terms of the input signals, $c_1(t)$ and $c_2(t)$.

---

* This report deals with the continuous (as opposed to sample-data) case of correlation. An analogous theory of correlation can be developed using discrete sequences and summations in place of the continuous functions and integrations.

** From the standpoint of the derivation which follows, it is unnecessary to fully define the averaging operation in equation (81). It must possess the property of allowing interchange of order with the integrations which are performed in the derivation. One possible definition that can be used is

$$\langle c_1(t)\, c_2(t-\tau) \rangle \equiv \lim_{n\to\infty} \sum_{k=1}^{n} \frac{{}_\kappa c_1(t)\, {}_\kappa c_2(t-\tau)}{n} \quad ; \tau \geq 0$$

when the right-hand side exists.

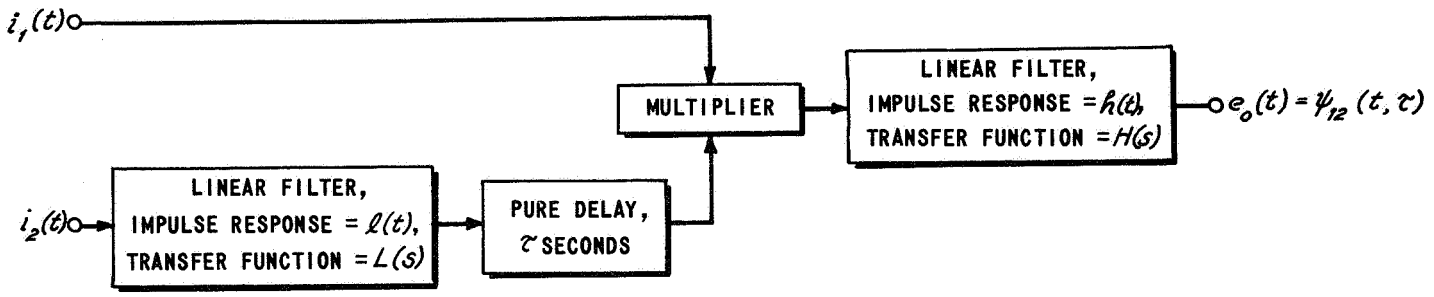


Figure 12 A CORRELATOR CONFIGURATION WHICH POSSESSES SPECIAL MATHEMATICAL PROPERTIES

If $h(t)$ and $\ell(t)$ represent physically realizable filtering operations, then the output signal is

$$e_o(t) \equiv \psi_{i,2}(t,\tau) = \int_0^\infty \int_0^\infty h(\lambda)\,\ell(\gamma)\,x_1(t-\lambda)\,x_2(t-\lambda-\gamma-\tau)\,d\gamma\,d\lambda \quad ; \tau \geq 0 \tag{82}$$

where λ and γ are variables of integration. If an ensemble average is performed over the output signals for each pair of input signals, then

$$\langle \psi_{i,2}(t,\tau) \rangle = \int_0^\infty \int_0^\infty h(\lambda)\,\ell(\gamma)\,\phi_{i,2}(t-\lambda,\ \gamma+\tau)\,d\gamma\,d\lambda \quad ; \tau \geq 0 \tag{83}$$

The independence property is exhibited by the integral of equation (83), wherein $h(t)$ operates only on the t axis and $\ell(t)$ operates only on the τ axis of the function $\phi_{i,2}(t,\tau)$. Since each filtering operation processes information in only one axis, the mathematical synthesis procedure for the **two** filters is simplified without loss of generality.

This analyzer configuration possesses other desirable properties. An important one is that the ensemble average output is unbiased for uncorrelated stationary or nonstationary input signals.* Additionally, the output of the correlator operating on $x_1(t)$ and $x_2(t)$ is unbiased over time if $x_1(t)$ and $x_2(t)$ are stationary, ergodic, and uncorrelated. These properties are important, for they show that the correlator does not yield steady, false indications of correlation.

Development of a Performance Measure

There are three sources of error which must be considered in the synthesis of a correlator for nonstationary signals. These sources of error exist regardless of the approach taken to nonstationary correlation analysis, even though they have not been fully recognized in the past. In order to undertake synthesis, it is necessary to develop a performance measure which accurately assesses these three sources of error. The measure may then be minimized so as to yield the optimal values of the filters, $h(t)$ and $\ell(t)$.

---

* An ensemble of uncorrelated nonstationary signals is defined to possess the property

$$\phi_{i,2}(t,\tau) \equiv \phi_{i,2}(t)\,\delta(\tau)$$

where $\delta(\tau)$ is the Dirac delta function.[27] Then, as long as $\tau > 0$, the analyzer configuration yields unbiased outputs.

The evaluation of the performance measure must not require a priori knowledge of the correlation function which is being detected, since the existence of this knowledge would make the subsequent correlation analysis pointless. One way of avoiding the problem of required a priori knowledge is to choose representative, known, test correlation functions and then to minimize the three sources of error in detecting them. It is then assumed that the correlation analyzer will work well in detecting correlation functions in the same general class as the known functions upon which it was optimized.

The optimal solution which is obtained will be heavily dependent upon the particular performance measure chosen to represent the errors. It is mathematically feasible to obtain a number of solutions by minimizing various performance measures, and "then to choose that particular solution which is best suited to the experimental data to be analyzed. In the next two sections, a single case is described in detail. Appendix I then summarizes two other representative cases.

The three sources of error in the output of a correlation analyzer are: 1) distortion (or error) of the true ensemble correlation function $\phi_{12}(t,\tau)$ along the A axis, 2) distortion (or error) of $\phi_{12}(t,\tau)$ along the τ axis, and 3) the inevitable noise or instability in the output of the analyzer which results from the components in the product $x_1(t)\,x_2(t-\tau)$ which are extraneous to the ensemble average, $\phi_{12}(t,\tau)$.* These three sources of error are most readily visualized by considering the product $x_1(t)\,x_2(t-\tau)$ to be made-up of 'the sum of the ensemble function $\phi_{12}(t,\tau)$ and the noiselike extraneous component $n(t,\tau)$. Then the output of the correlator is

$$\Psi_{12}(t,\tau) = \int_0^{\infty}\int_0^{\infty} h(\lambda)\,l(\gamma)\,\phi_{12}(t-\lambda,\gamma+\tau)\,d\gamma\,d\lambda + \int_0^{\infty}\int_0^{\infty} h(\lambda)\,l(\gamma)\,n(t-\lambda,\gamma+\tau)\,d\gamma\,d\lambda \quad ; \quad \tau \geq 0 \qquad (84)$$

The objective will be to make the first integral approximate $\phi_{12}(t,\tau)$ as closely as possible while making the second integral as small in value as possible.

Distortion Along the A Axis. - In order to quantify the distortion of the true ensemble correlation function produced by the output of the correlation analyzer, the first integral of equation (84) is examined. -An error representing the distortion can be defined as the difference between the true correlation function and this first integral, that is

$$e_a(t,\tau) \equiv \phi_{12}(t,\tau) - \int_0^{\infty}\int_0^{\infty} h(\lambda)\,l(\gamma)\,\phi_{12}(t-\lambda,\gamma+\tau)\,d\gamma\,d\lambda \quad ; \quad \tau \geq 0 \qquad (85)$$

---

* The third source of error may be equivalently considered as the result of finite statistical averaging time, which always produces instabilities in the statistical estimates obtained.

At this point it is necessary to choose an ensemble correlation function and an error measure which will serve as a test of the distortion along the t axis. A rather elementary, but effective, test function is the following:

$$\phi_{i_2}(t,\tau) = \begin{cases} t \; ; & t \geq 0 \\ 0 \; ; & t < 0 \end{cases} \quad ; \; \tau \geq 0 \qquad (86)$$

That is, the test correlation function is a ramp beginning at $t = 0$. The correlation function is assumed uniform in τ for $\tau \geq 0$.* (This ensemble correlation function can be realized by ideal components as shown in Figure 13.) An error measure can be formed by squaring each side of equation (85) and integrating over time. Then it is found upon substitution that

$$P_t \equiv \int_0^\infty e_d^2(t,\tau)\, dt = \int_0^\infty \left[t - L\, r_h(t)\right]^2 dt \qquad (87)$$

where $L \equiv \int_0^\infty l(\tau)\, d\tau$

and $r_h(t) \equiv \int_0^t h(\lambda)(t-\lambda)\, d\lambda$

= ramp response of the filter whose impulse response is $h(t)$.

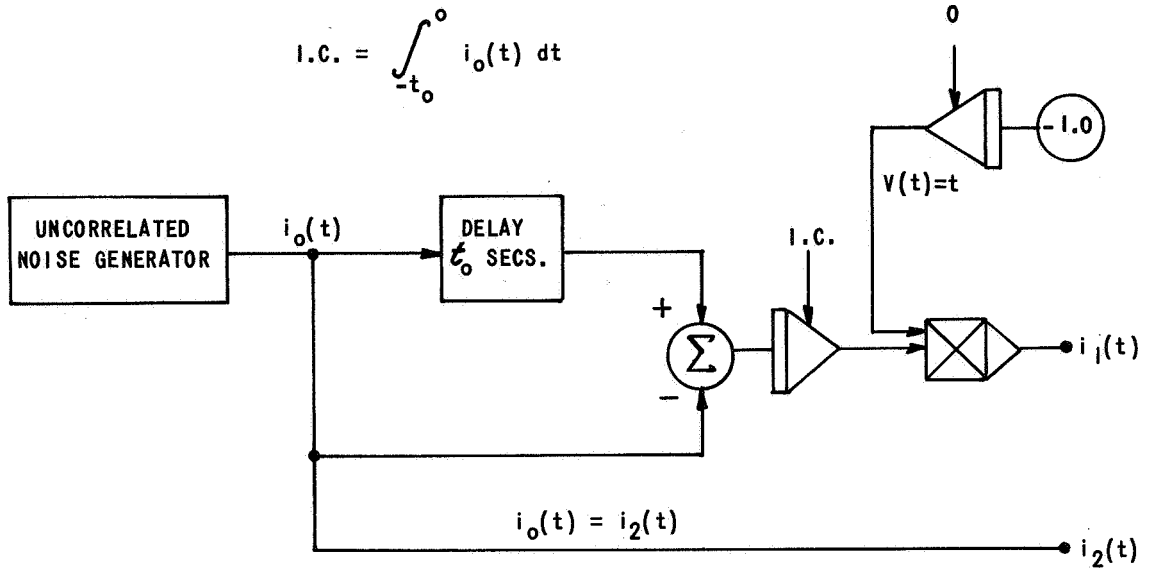
The quantity P_t can be considered as a general measure of the analyzer's distortion of $\phi_{i_2}(t,\tau)$ along the t axis.

Distortion Along the τ Axis. - Distortion along the τ axis can be handled in similar, but slightly more complicated, manner, Equation (85) is used once again, but a different test function and measure are considered. One test function which adequately tests the distortion along the τ axis is the triangular function. In practice, nearly triangular correlation functions are often encountered, since in general correlations tend to zero as τ becomes large. It can be assumed that, if the distortion is relatively small in detecting a triangular correlation function, the distortion will also be small for any correlation function which can be approximated by a group of staggered triangular functions. Let

$$\phi_{i_2}(t,\tau) = \begin{cases} \tau_0 - \tau \; ; & 0 \leq \tau \leq \tau_0 \\ 0 \; ; & \tau > \tau_0 \end{cases} \quad ; \; \tau \geq 0 \qquad (88)$$

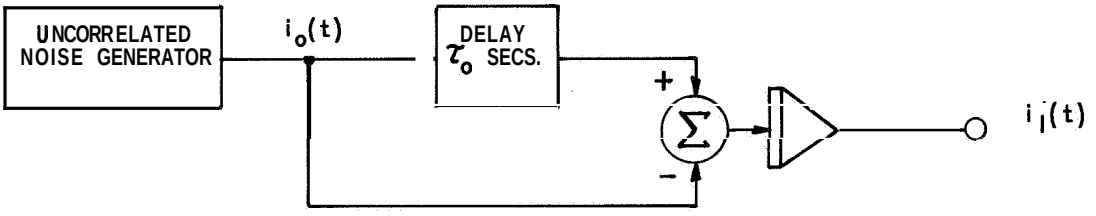
---

* The assumption that the test function is uniform in τ makes possible the study of distortion along the t axis without interaction of the τ axis distortion.



$$\phi_{i2}(t,\tau) = \begin{cases} t\ ;\ 0 \le \tau \le t_o \text{ AND } t \ge 0 \\ 0\ ;\ \tau > t_o \text{ AND, OR } t < 0 \end{cases}$$

Figure 13 IDEAL COMPONENT CONFIGURATION FOR GENERATING AN ENSEMBLE CORRELATION FUNCTION WHICH IS A RAMP IN t.



$$\phi_{11}(t,\tau) = \begin{cases} \tau_o - \tau\ ;\ 0 \le \tau \le \tau_o \\ 0\ ;\ \tau \ge \tau_o \end{cases}$$

Figure 14 IDEAL COMPONENT CONFIGURATION FOR GENERATING AN ENSEMBLE CORRELATION FUNCTION WHICH IS TRIANGULAR IN τ.

In this case, it is being assumed that the ensemble correlation function is uniform in time; that is, it is stationary. (This ensemble correlation function can be realized by ideal components as shown in Figure (14).) If the error described by equation (87) is squared and integrated over γ, the result will serve as a measure of distortion along the γ axis of $\phi_2(t,\gamma)$:

$$P_\tau \equiv \int_0^\infty e_d^2(t,\tau)\,d\tau = \int_0^{\tau_0} \left[\tau_0 - \tau - \int_0^\infty h(\lambda)\,d\lambda \cdot \int_0^{\tau_0-\tau} l(\gamma)(\tau_0-\tau-\gamma)\,d\gamma\right]^2 d\tau \qquad (89)$$

$$= \int_0^{\tau_0} \left[\mu - H\eta_l(\mu)\right]^2 d\mu$$

where $H \equiv \int_0^\infty h(\lambda)\,d\lambda$

and $\eta_l(\mu) \equiv \int_0^\mu l(\gamma)(\mu-\gamma)\,d\gamma$

= ramp response of the filter whose impulse response is $l(\mu)$.

Noise in the Output of the Correlator. - A correlator of nonstationary signals, which operates upon a member pair of an ensemble of functions, will have a noise-like component in its output signal. This noise or instability component is the result of that portion of the input signal pair which is unrelated to the true ensemble correlation function. The second integral of equation (84) describes the effect of this noise on the detected correlation function.

As discussed previously, in order to assess the effect of this noise component, it is necessary to choose a noise signal and a measure which adequately test the smoothing characteristics of the analyzer. It must be remembered, however, that an average other than an ensemble average must be used for the noise test, since an ensemble average of the function $n(t,\gamma)$ is zero. Essentially, a single pair of noisy time waveforms must be used to test the immunity of the analyzer to noise or instability. A very good test is the response of the analyzer to uncorrelated signals which are generated by a stationary and ergodic Gaussian process. This set of input signals can be considered as an extreme test because signal power exists at all frequencies, and therefore all frequencies are tested. A good measure of the noise or

instability is the variance of the output of the correlator for the two uncorrelated input signals described above. This variance can be written as

$$P_n \equiv \overline{e_o^2(t)} = \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty h(\lambda_1)\,h(\lambda_2)\,l(\gamma_1)\,l(\gamma_2) \tag{90}$$

$$\times\, \overline{n_1(t-\lambda_1)\,n_1(t-\lambda_2)\,n_2(t-\lambda_1-\gamma_1-\tau)\,n_2(t-\lambda_2-\gamma_2-\tau)}$$

$$\times\, d\gamma_1\, d\gamma_2\, d\lambda_1\, d\lambda_2 \;;\quad \tau > 0$$

where a bar over a quantity indicates a time average.* If advantage is taken of the theorems for Gaussian processes,[28] it is found that

$$\overline{n_1(t-\lambda_1)\,n_1(t-\lambda_2)\,n_2(t-\lambda_1-\gamma_1-\tau)\,n_2(t-\lambda_2-\gamma_2-\tau)} \;=$$

$$\delta(\lambda_2-\lambda_1)\cdot\delta(\lambda_2-\lambda_1+\gamma_2-\gamma_1) \tag{91}$$

$$+\,\delta(\gamma_1+\tau)\cdot\delta(\gamma_2+\tau)$$

$$+\,\delta(\lambda_2-\lambda_1+\gamma_2+\tau)\cdot\delta(\lambda_1-\lambda_2+\gamma_1+\tau)$$

If equation (91) is substituted into equation (90), it is easily shown that the second and third terms of equation (91) contribute nothing to the value of the variance for $\tau > 0$. Then, integration of the remaining term first with respect to λ_1 and then with respect to γ_1 yields the result that

$$P_n = \int_0^\infty h^2(\lambda)\,d\lambda \cdot \int_0^\infty l^2(\gamma)\,d\gamma \;;\quad \tau > 0 \tag{92}$$

Equation (92) describes in a straightforward way the variance of the correlator's output for an uncorrelated set of input signals. It represents, in a general way, the noise or instability in the output of the correlator.

---

* The variance of the output is equal to the mean-squared value of the output, since for uncorrelated input signals the output is unbiased.

Performance Measure. - The performance measure for the correlation analyzer may now be defined as the sum of the three error measures given by equations (87), (89), and (92). However, it must be recognized that the relative weights to be placed upon each measure are arbitrary; that is, it is a matter of choice to decide how important each of the three sources of error is to the design of the analyzer. Accordingly, the performance measure should be written in the form

$$\Theta = \lambda_t^2 P_t + \lambda_\gamma^2 P_\gamma + P_n \tag{93}$$

where λ_t^2 and λ_γ^2 are arbitrary positive constants. There is no need to associate a coefficient with the measure P_n, because Θ may always be normalized such that the coefficient could be made equal to unity, without any change in the optimal uesign.

It is important to realize that there are two uncertainties with which one must deal in the design of a nonstationary correlation (or spectral) analyzer. These uncertainties make it impossible to detect a statistical nonstationary function, such as a correlation function without error. The two constants λ_t^2 and λ_γ^2 represent the compromises which must be reached between errors on each axis of the detected correlation functions and noise or instability in the output of the analyzer. These uncertainties are fundamental and must be recognized if a meaningful correlator design is to be developed.

Design of the Correlation Analyzer

Since the nonstationary correlation analyzer configuration has already been specified, the completion of the analyzer design consists of determining the two filtering operations represented by $h(t)$ and $e(t)$, as shown in Figure 12. These filters can be determined by minimization of the performance measure, Θ, given by equation (93). Minimization of this measure will result in the optimum analyzer design, thereby specifying $h(t)$ and $e(t)$.

In the minimization process, the following assumptions will be made regarding the filtering operations:

$$L = \int_0^\infty e(\tau)\, d\tau \equiv 1.0 \tag{94}$$

$$H = \int_0^\infty h(\lambda)\, d\lambda \equiv 1.0 \tag{95}$$

$$r_e(0) = \dot{r}_e(0) = 0 \tag{96}$$

$$r_h(0) = \dot{r}_h(0) = 0 \tag{97}$$

$$\lim_{\gamma \to \infty} [\gamma - r_z(\gamma)] = 0 \tag{98}$$

and

$$\lim_{\lambda \to \infty} [\lambda - r_h(\lambda)] = 0 \tag{99}$$

The first two conditions require that the unit step response of each filter will eventually exhibit correspondence by settling at unity. The second **two** conditions force the ramp response and the step response of each filter to begin at zero amplitude. The last **two** conditions require that the unit ramp response of each filter will eventually exhibit correspondence by settling to a unit ramp. Under this group of assumptions, the performance measure may be written in the form

$$\Theta = \lambda_t^2 \int_0^\infty [\lambda - r_h(\lambda)]^2 d\lambda + \lambda_\gamma^2 \int_0^{\gamma_0} [\gamma - r_z(\gamma)]^2 d\gamma + \int_0^\infty \ddot{r}_z^2(\gamma) d\gamma \cdot \int_0^\infty \ddot{r}_h^2(\lambda) d\lambda \tag{100}$$

where the impulse response of each filter is written as the second derivative of the corresponding ramp response.

If Θ in equation (100) is minimized, then it will be found that one condition on the optimal solution is that $\ell(\gamma) = 0$ for all $\gamma > \gamma_0$. This condition is the result of the finite upper limit on the integral in P_γ. The requirement that $\ell(\gamma)$ be zero over a finite interval will require that the transfer function corresponding to $\ell(\gamma)$ be other than the ratio of **two** finite polynomials in the complex frequency variable. Accordingly, a spectral approximation procedure would have to .be developed, and would yield **only** an approximate optimal solution. It is believed that the spectral approximation procedure would unduly complicate the solution to this problem, and that it would not reduce by any worthwhile amount the errors in the correlation analyzer.

A straightforward method for circumventing the above problem is to add a further condition to the performance measure which insures competition between the ramp response error and the noise error of $\ell(\gamma)$ for all positive γ. The condition which should be added is

$$\lambda_\gamma^2 \int_{\gamma_0}^{\bar{\ }} [\gamma - r_z(\gamma)]^2 d\gamma$$

which produces a form of the performance measure which yields to analysis and which produces strictly realizable transfer functions for $\ell(\tau)$ and $h(\lambda)$. The new performance measure is the same as that given in equation (100) except that the upper limit on the second integral is now taken to be infinite.

Calculus-of-Variations Solution. - The performance measure can be minimized by means of the calculus of variations. The method of solution used herein is similar to (but is more complex than) one reported previously in regard to optimization of radar trackers.[29] The solution is initiated by taking variations in each of the functions $r_\ell(\tau)$ and $r_h(\lambda)$. These variations, when substituted into the performance measure, O , yield the following two variational equations:

$$I(\alpha_\ell) = \int_0^\infty \left[\ddot{r}_\ell(\tau) + \alpha_\ell \ddot{n}_\ell(\tau)\right]^2 d\tau \cdot \int_0^\infty \ddot{r}_h^{\,2}(\lambda)\, d\lambda + \lambda_t^2 \int_0^\infty \left[\lambda - r_h(\lambda)\right]^2 d\lambda + \lambda_\tau^2 \int_0^\infty \left[\tau - r_\ell(\tau) - \alpha_\ell n_\ell(\tau)\right]^2 d\tau \tag{101}$$

and

$$I(\alpha_h) = \int_0^\infty \left[\ddot{r}_h(\lambda) + \alpha_h \ddot{n}_h(\lambda)\right]^2 d\lambda \cdot \int_0^\infty \ddot{r}_\ell^{\,2}(\tau)\, d\tau + \lambda_\tau^2 \int_0^\infty \left[\tau - r_\ell(\tau)\right]^2 d\tau + \lambda_t^2 \int_0^\infty \left[\lambda - r_h(\lambda) - \alpha_h n_h(\lambda)\right]^2 d\lambda \tag{102}$$

where α_ℓ and α_h are small, arbitrary constants and $n_\ell(\tau)$ and $n_h(\lambda)$ are the variations in $r_\ell(\tau)$ and $r_h(\lambda)$, respectively. From this point the derivation is carried forward for the optimal value of $\ell(\tau)$ only. The form obtained for $h(\lambda)$ will be the same as that obtained for $\ell(\tau)$.

The usual calculus-of-variations procedure is followed. Form the equation

$$\left.\frac{\partial I(\alpha_\ell)}{\partial \alpha_\ell}\right|_{\alpha_\ell = 0} = 0 \tag{103}$$

Integration by parts, applied twice, then allows the resulting equation to be written as

$$\left[\ddot{r}_\ell(\gamma)\,\dot{\eta}_\ell(\gamma)\Big|_0^\infty - \dddot{r}_\ell(\gamma)\,\eta_\ell(\gamma)\Big|_0^\infty + \int_0^\infty \dddot{r}_\ell(\gamma)\,\eta_\ell(\gamma)\,d\gamma\right]\int_0^\infty \ddot{r}_k^{\,2}(\lambda)\,d\lambda$$

$$-\lambda_\gamma^2 \int_0^\infty [\gamma - r_\ell(\gamma)]\,\eta_\ell(\gamma)\,d\gamma = 0 \qquad (104)$$

The first two terms of this equation are each equal to zero. The first term is zero because the step response of $\ell(\gamma)$ has been specified at both $\gamma = 0$ [equation (97)] and $\gamma = \infty$ [equation (94) which requires that $\lim_{\gamma\to\infty} \ell(\gamma) = 0$]. Therefore, the derivative of the variation $\dot{\eta}_\ell$ must be zero at these points, which makes the first term equal to zero. The second term is zero because $r_\ell(\gamma)$ has been specified at $\gamma = 0$ and $\gamma = \infty$, thereby making the, variation η_ℓ equal to zero at these points. By the fundamental theorem of the calculus of variations, the Euler differential equation for the extremals is obtained:

$$\ddddot{r}_\ell(\gamma) = 4k_\gamma^4[\gamma - r_\ell(\gamma)] \qquad (105)$$

where

$$k_\gamma \equiv \left|\frac{1}{\sqrt{2}}\left\{\frac{\lambda_\gamma^2}{\left[\int_0^\infty k^2(\lambda)\,d\lambda\right]}\right\}^{1/4}\right| \qquad (106)$$

The solution of this differential equation is of the form

$$r_\ell(\gamma) = \gamma + Ae^{k_\gamma[-1+j]\gamma} + Be^{k_\gamma[-1-j]\gamma} + Ce^{k_\gamma[1+j]\gamma} + De^{k_\gamma[1-j]\gamma} \qquad (107)$$

The solution is valid for $\gamma \geq 0$ only. Since the solution must eventually converge to a unit ramp, the growing exponential terms must have zero coefficients. Consequently, $C = 0$ and $D = 0$. Also, if advantage is taken of the specifications given by equation (96), the solution can be shown to take the form

$$r_\ell(\gamma) = \gamma - \frac{1}{j2k_\gamma}e^{k_\gamma[-1+j]\gamma} + \frac{1}{j2k_\gamma}e^{k_\gamma[-1-j]\gamma} \qquad (108)$$

$$= \gamma - \frac{1}{k_\gamma}e^{-k_\gamma\gamma}\sin k_\gamma\gamma$$

The Laplace transform of this ramp response, $R_L(s)$, is found to be

$$R_L(s) = \frac{2k_\tau s + 2k_\tau^2}{s^2(s^2 + 2k_\tau s + 2k_\tau^2)} \tag{109}$$

Therefore, the transfer function of the optimal filter, $L(s)$, (the Laplace transform of the impulse response $l(\tau)$), is given by

$$L(s) = \frac{2k_\tau s + 2k_\tau^2}{s^2 + 2k_\tau s + 2k_\tau^2} \tag{110}$$

The transfer function, $H(s)$, of the optimal filter whose impulse response is $h(\lambda)$ is found to possess the same form:

$$H(s) = \frac{2k_t s + 2k_t^2}{s^2 + 2k_t s + 2k_t^2} \tag{111}$$

where

$$k_t = \left| \frac{1}{\sqrt{2}} \left\{ \frac{\lambda_t^2}{\left[\int_0^\infty l^2(\tau)\, d\tau \right]} \right\}^{1/4} \right| \tag{112}$$

It is seen that the constants k_τ and k_t are interdependent. Consequently, if k_τ and k_t are to be evaluated in terms of λ_τ^2 and λ_t^2 (the original weights of the performance measure), it is necessary to remove these dependencies. This can be done by substituting the optimal values of the $l(\tau)$ and $h(\lambda)$ into equations (106) and (112). Simultaneous solution then yields

$$k_\tau = \frac{1}{6^{1/5}} \left(\frac{\lambda_\tau^4}{\lambda_t} \right)^{2/15} \tag{113}$$

and

$$k_t = \frac{1}{6^{1/5}} \left(\frac{\lambda_t^4}{\lambda_\tau} \right)^{2/15} \tag{114}$$

Accordingly, equations (110), (111), (113), and (114) specify the forms of the optimal filters in terms of the weights in the performance measure.* In certain cases, the constants k_τ and k_t may be chosen first, in which

---

* This calculus-of-variations solution yields only one extremal. It is easy to prove that this extremal produces minimum correlator error (according to the given performance measure). The proof is obtained by showing that $I(a_l)$ in equation (101) and $I(a_h)$ in equation (102) are always greater than or equal to Θ when the optimal values of $n_l(\tau)$ and $n_h(\lambda)$ are substituted.

case it may be desirable to determine the values of λ_γ^2 and λ_t^2, which have then been implicitly chosen. The reciprocal relationships are found to be:

$$\lambda_\gamma^2 = 6 k_t k_\gamma^4 \tag{115}$$

and

$$\lambda_t^2 = 6 k_\gamma k_t^4 \tag{116}$$

The design of the nonstationary correlator has thus been completed.

It should be mentioned that, although it is possible to evaluate the three terms in the performance measure for any setting of λ_γ and λ_t, the results are of little practical value in determining good settings for λ_γ and λ_t. Experimental study has shown that a good procedure is to set k_γ based upon bandwidth and resolution considerations of $L(s)$. Afterward, k_t is set by means of a preliminary ekperiment so that sufficient noise smoothing is obtained.

Experimental Study

An extensive experimental study was performed to verify and determine the usefulness of the theory of correlation described in this report. The theory as described herein is exact, so that it is only necessary to insure its correctness. The correctness was established by having the analyzer operate on signals which were similar to the three test signals used in the mathematical optimization process. The usefulness of the theory was studied by having the analyzer detect a known nonstationary cross-correlation function which is similar to that obtained in practice, but is entirely different from the three test functions. The experiments and results are briefly described in this section.

The experiments were all performed on an E. A. I. model TR-48 analog computer with an additional TR-20 used as a slave unit. A high quality wide-band noise source was used to approximate the uncorrelated input signal required in each experiment. This signal was recorded on one channel of a 7 channel F.M. tape recorder. Then, since each channel had the record and playback heads physically separated, it was possible to obtain pure time delay of the original noise signal by recording the output of one channel onto the next channel. Special filtering and amplitude reference signals were used to maintain the quality of each reproduction, In this way it was possible to obtain six equally spaced values of delay of the original waveform (the samples being 0.733 seconds apart). In order to obtain delay values between those available from the tape recorder, a new technique was used which requires only standard analog computer components.[30] Thus, the tape recorder was used for the large delays, and the new delay technique was used for additional vernier delays

between those of the tape recorder. All of the experimental runs were made serially; that is, data for each given value of delay were recorded separately. These data were then combined on the three dimensional graphs, which are the results of the experiments.

Experiment 1

This experiment was composed of three parts, each part representing one term of the performance measure, Θ . The objective was to determine in a qualitative way the ability of the analyzer to detect the correlation functions of signals similar to those for which it had been optimized. The three parts of the experiment and the results obtained are briefly described below:

1) Autocorrelation of the wide-band, stationary, random input signal. This test approximates the theoretical uncorrelated noise test of the analyzer. The steady-state results of the experiment (See Figure 15) indicate that the input signals are uncorrelated for the three values of delay which were used. They also show the inevitable noise or instability which always accompanies the detection of a random process using finite averaging time. The line of data for $\gamma = 0$ is not infinite in amplitude (it is infinite for a theoretical uncorrelated process) because of the finite bandwidth of the input signal and because of the smoothing of the γ axis information by the filter, $\ell(t)$. In this experiment $k_\gamma = 1.06$ and $k_\ell = 0.00524$.

2) Cross-correlation of the output of a network, whose impulse response is a single exponential with a time constant of 2.2 seconds, with the wide-band input to the network. This test approximates the distortion test involving the stationary, triangular correlation function. It would be expected that the errors in detecting this correlation function should be small, since the analyzer was optimized on a similar correlation function.

   The steady-state results of the experiment (See Figure 16) show that the errors in detection are indeed small. The theoretical curve of the correlation function is the dotted line at the front of the plot. It can be concluded from this plot that the analyzer is responding in a manner which is consistent with the mathematical results obtained earlier. The values of k_γ and k_ℓ were the same as for the above experiment.

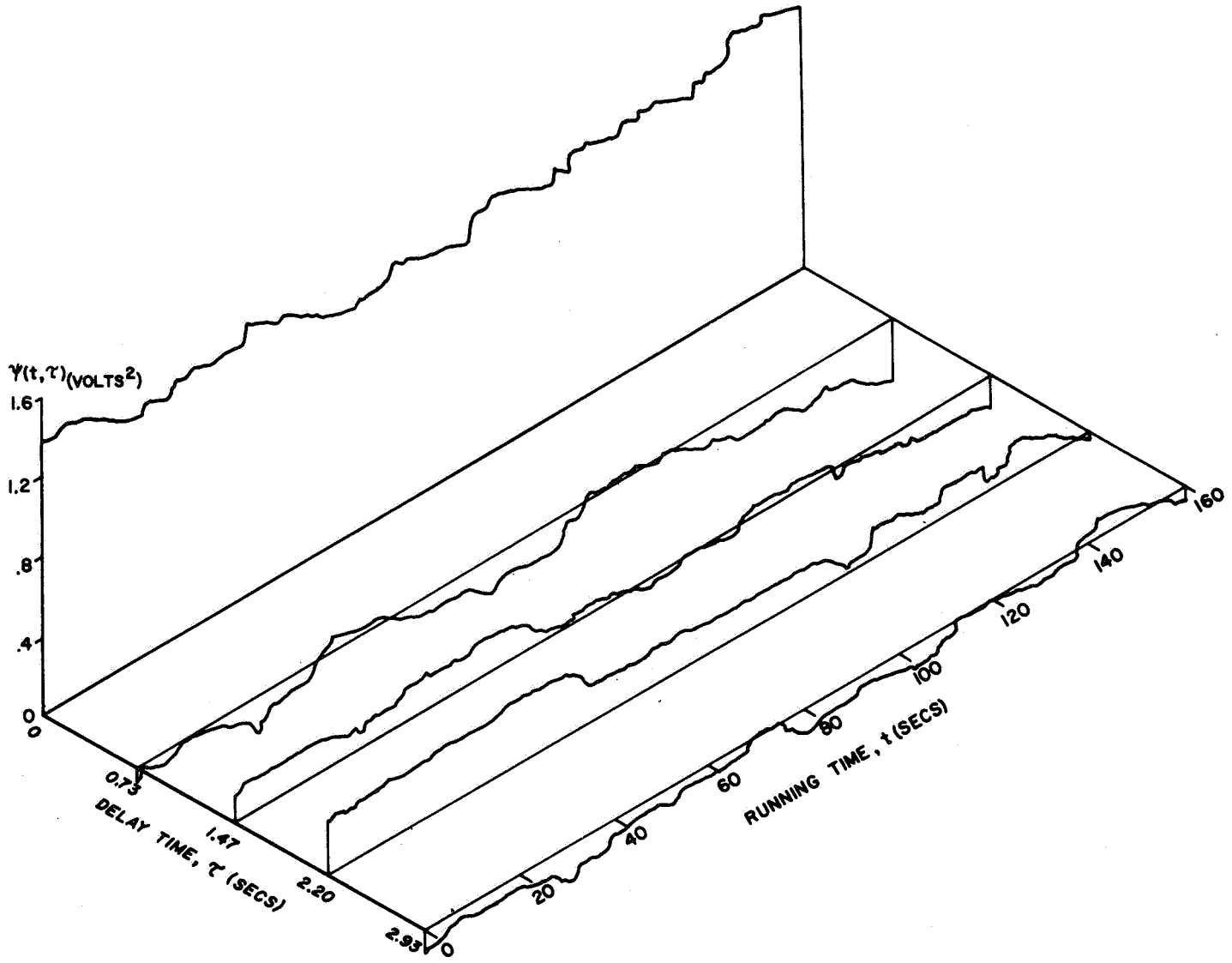


Figure 15 EXPERIMENTAL STEADY-STATE RESPONSE OF THE CORRELATOR TO A WIDE-BAND NOISE INPUT SIGNAL

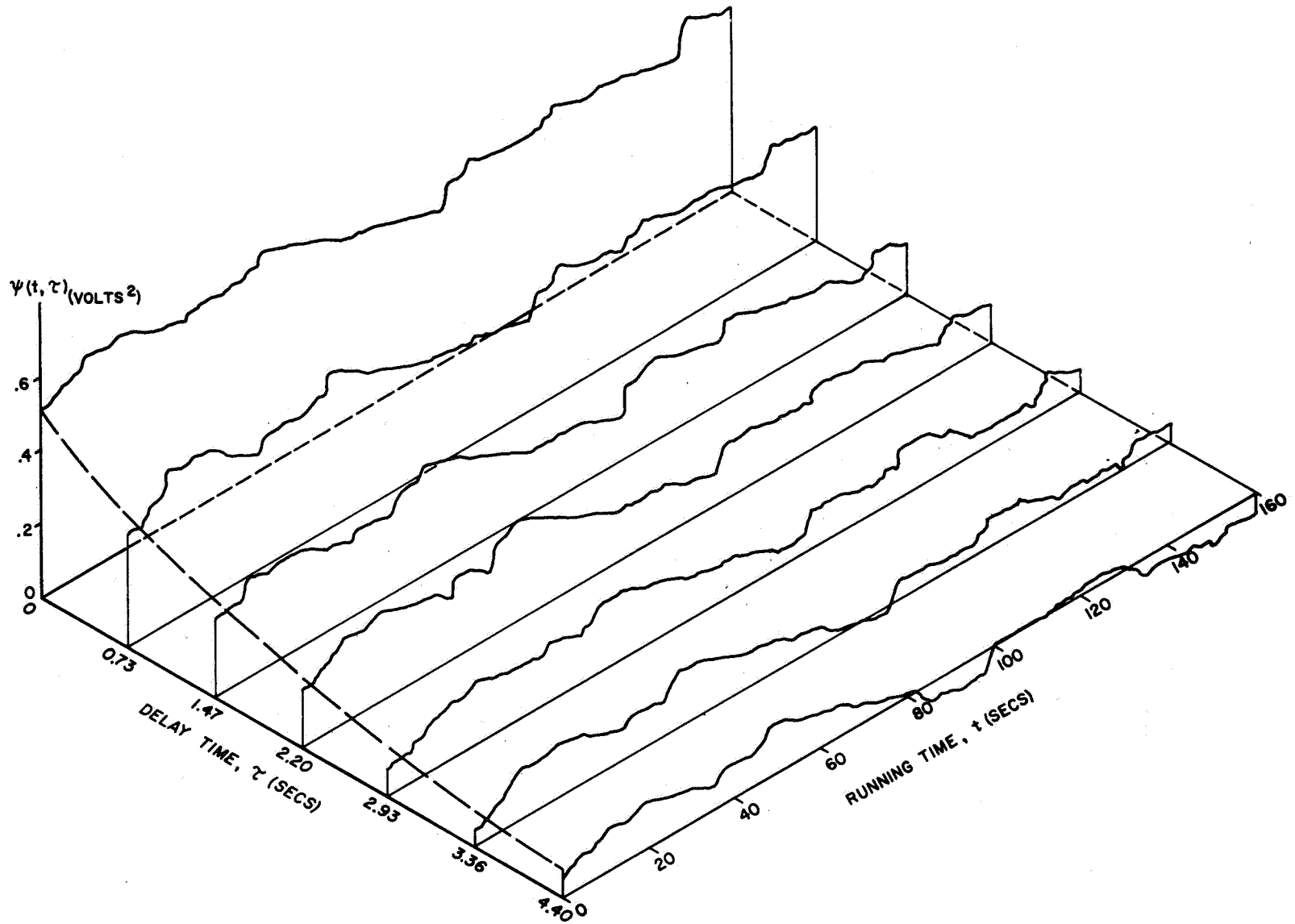


Figure 16 EXPERIMENTAL STEADY-STATE RESPONSE OF THE CORRELATOR TO A SIGNAL WHOSE ENSEMBLE CORRELATION FUNCTION IS AN EXPONENTIAL IN τ.

3) Cross-correlation of nonstationary signals whose theoretical ensemble cross-correlation function is a slowly rising ramp in time. This test is similar to the theoretical test of distortion along the τ axis. An experimental generator similar to that shown in Figure 13 was used. The analyzer's transient output beginning at $t = 0$ was recorded. Because of the similarity of the responses, only one has been plotted.

The results of this experiment (See Figure 17) show that the analyzer is capable of accurately follpwing slow ramp changes in the nonstationary correlation function. It shows neither appreciable undershoot nor overshoot. Therefore, it is quite clear that the analyzer is responding in an optimum manner for this form of input. Once again the same values of k_1 and k_2 were used.

The results of Experiment 1 show conclusively that the analyzer responds in practice in a manner which the foregoing theory has indicated.

Experiment 2

The objective of the second experiment was to evaluate the analyzer's performance in detecting a nonstationary correlation function which was significantly different from the three upon which it had been optimized, but is typical of correlation functions found in practical situations.

The nonstationary input signals for the analyzer were generated by applying the wide-band input signal to the inputs of each of two networks. After the analyzer had reached a steady-state condition in measuring the cross-correlation between output and input of the first network, the analyzer input was switched so as to measure the cross-correlation between the output and input of the second network. This type of arrangement produces a step change in time in the ensemble correlation function being detected. Also, since both networks are known, it makes possible the study of distortion along the τ axis.

The two networks used to generate the input signals had impulse responses given by

$$g_1(t) = -3.41\, e^{-\frac{t}{0.75}} + 1.19\, e^{-\frac{t}{2.0}} \qquad (117)$$

and

$$g_2(t) = 2.10\, e^{-0.6t} \sin 1.91\, t \qquad (118)$$

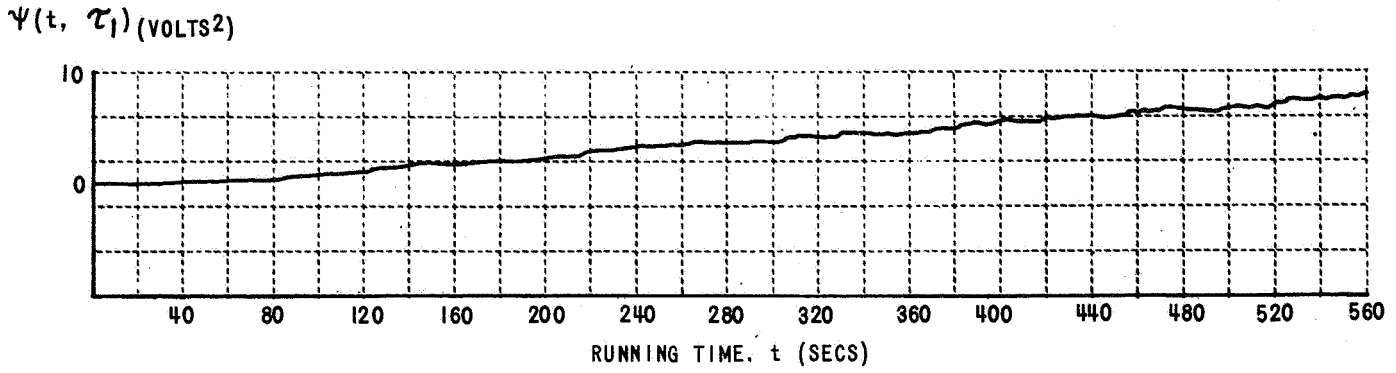


Figure 17 EXPERIMENTAL RESPONSE OF THE CORRELATOR TO A SIGNAL WHOSE ENSEMBLE CORRELATION FUNCTION IS AN EXPONENTIAL IN t

The τ axis distortion that the analyzer produces (in steady-state) in detecting cross-correlation functions of this form can be predicted by performing an unfolded convolution of $\ell(t)$ with each of the above responses. (This result can be proved by obtaining the ensemble cross-correlation function between the output of the networks whose responses are $g_1(t)$ and $g_2(t)$ and the output of the analyzer network given by $\ell(t)$.) Thus, this distorted steady-state output of the analyzer should be

$$g_{01}(\tau) = \int_0^\infty \ell(\gamma)\, g_1(\gamma+\tau)\, d\gamma \qquad (119)$$

for the first network, and

$$g_{02}(\tau) = \int_0^\infty \ell(\gamma)\, g_2(\gamma+\tau)\, d\gamma \qquad (120)$$

for the second network. These two theortical, distorted responses were computed digitally and have been plotted using dotted lines at each end of the experimental plot of Figure 18 .

The remainder of the plot of Figure (18)shows the experimental data obtained. It is seen that the τ axis distortion in steady-state has been accurately predicted. The fact that the correlation function changes abruptly in time (at t = 160 seconds) rather than changing as a ramp, indicates that $h(t)$ is not longer optimum. As a result, the response of the analyzer overshoots somewhat before settling to the new values of the correlation function. Once again the noise or instability is present. The analyzer constants were set at $k_\tau = 2.12$ and $k_t = 0.00524$. This setting of constants afforded less noise smoothing than was used in the first set of experiments.

Although distortion exists along each axis of the correlation function and noise exists in the measurements, it is evident that the detected information would be valuable in classifying and understanding the process from which the signals were generated. The ability to detect nonstationary correlation functions using the techniques developed in this report will be heavily dependent upon the degree of resolution required along the τ axis and the degree of dependence of the ensemble correlation function upon its past values (in t). These two factors determine the amount of noise smoothing which may be incorporated in the analyzer without producing excessive distortion in the detected information.

Conclusions to the Study of Nonstationary Signal Processing

It has been shown that a correlation analyzer for nonstationary signals may be synthesized by proper definition of a nonstationary correlation function, choice of an analyzer configuration, and choice of a performance measure. Subsequently, the optimal filtering operations may be determined by means of the calculus of variations.

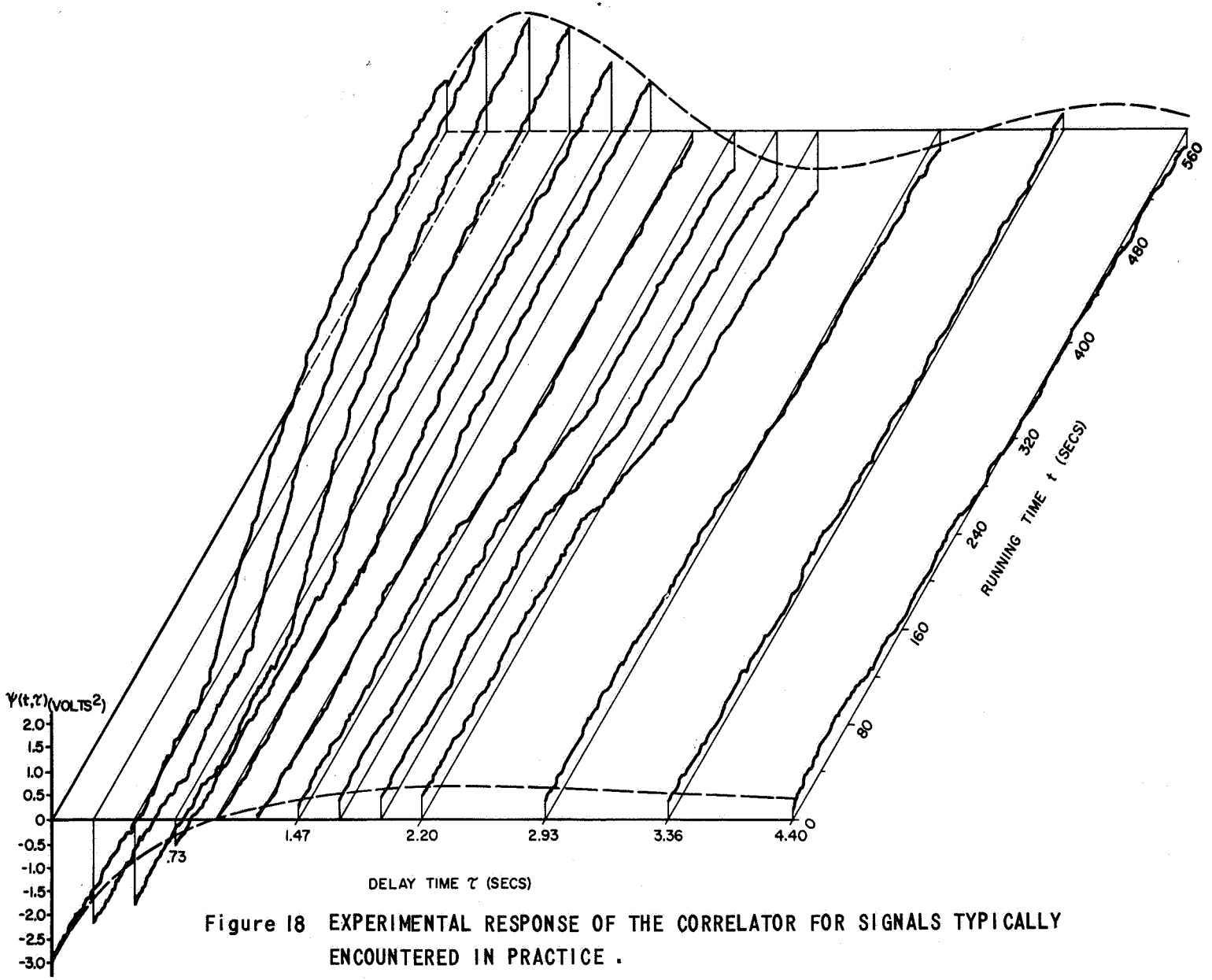


Figure 18 EXPERIMENTAL RESPONSE OF THE CORRELATOR FOR SIGNALS TYPICALLY ENCOUNTERED IN PRACTICE .

Two fundamental uncertainties exist in the design of a correlator for nonstationary signals. The first involves a compromise between distortion along the running time axis of the detected correlation function and noise in the output of the analyzer. The second involves a similar compromise between distortion along the delay (or τ) axis of the detected correlation function and noise in the output of the analyzer. The detrimental effects of uncertainties can be minimized by performing a minimization of a performance measure which accurately assesses these detrimental effects. The analyzer configuration which results is then considered to be optimum.

It is anticipated that better results could be obtained with this new nonstationary analyzer as compared with other analyzers. The fact that other analyzers do not include the filtering operation represented by $\ell(t)$ indicates that maximum noise smoothing is not being accomplished. Moreover, other analyzers are not generally optimized to take maximum advantage of the limited amount of raw data available.

Experiments have been performed which show that good correspondence exists between the mathematical aspects of the theory and their counterparts in practice. Moreover, for a typical example, but one for which the analyzer was not optimized, the analyzer continued to operate acceptably.

CONCLUSIONS

The purpose of the research presented in this report was to satisfy the Computation Laboratory requirements of (1) studying and applying the available random data processing techniques to the existing MSFC problems, (2) develop new and improved techniques of data processing. The following discussion indicates that the above requirements are satisfied.

Digital Filtering. - Selection of an appropriate sampling interval which produces negligible frequency folding is paramount to accurate digital data processing.

The vast amount of literature available which describes digital simulation of transfer functions from the time response point of view can be utilized to produce prewhitening filters having specific frequency characteristics.

Application of equation (13) along with the information contained in Appendix B allows the synthesis of discrete filters with sharp frequency cutoff characteristics. The filters so generated are applicable to all non-real-time data processing. The same techniques can be applied to obtain filters with integrating or differentiating qualities.

Taking the Tustin Transform (see Reference 15) of an analog notch filter will produce a digital filter which can be used for prewhitening, with the possibility of total rejection of one frequency. These notch filters contain relatively few weights.

In situations where the power spectral density function of only a band of frequencies is of interest, digital hetrodyning as discussed in the report may provide a computational time savings in data processing.

Correlation Functions. - After reading the analysis of different methods of estimating correlation function, one should conclude that modifications should be made to any existing computational technique that does not consider (1) the accuracy of estimates and (2) the computer time required. Many types of correlation function estimators are given (autocorrelation being a special case of cross-correlation). Extensive study of the "half-polarity" correlator is presented. Computer programs are outlined, which will calculate, in minimum time, the "half-polarity" and "full-precision" correlation functions. It is also suggested that correlation computational techniques given in the reference are applicable.

Optimal Smoothing of PSD. - The appropriate application of equation **(69)** will produce spectral estimates with greater accuracy and also eliminate the need for prewhitening of the signal prior to processing.

Deterministic Processing. - In many situations the source of the signal to be analyzed and the information to be extracted may dictate the use of a deterministic approach as indicated by equation (70). This approach requires no smoothing or prewhitening and will not yield negative power in the final spectrum.

Nonstationary Spectrum Analysis. - A new and essentially different approach to the spectrum analysis of nonstationary signals has been described. This technique utilizes both time and frequency smoothing to determine the presence of bands of frequencies. A priori knowledge of the time variations may enable the spectrum analyzer to be designed in an optimum manner for a particular type of nonstationary process. A test signal representing a wide class on nonstationary signals has been analyzed by this new technique and the results show that the method is indeed useful for determining the frequency content as well as the nature of the time variations for this class of nonstationary signals.

Nonstationary Correlation Functions. - It has been shown that a correlation analyzer for nonstationary signals may be synthesized by proper definition of a nonstationary correlation function, choice of an analyzer configuration, and choice of a performance measure. Subsequently, the optimal filtering operations may be determined by means of the calculus of variations.

Two fundamental uncertainties exist in the design of a correlator for nonstationary signals. The first involves a compromise between distortion along the running time axis of the detected correlation function and noise in the output of the analyzer. The second involves a similar compromise between distortion along the delay (or β) axis of the detected correlation function and noise in the output of the analyzer. The detrimental effects of uncertanties can be minimized by performing a minimization of a performance measure which accurately assesses these detrimental effects. The analyzer configuration which results is then considered to be optimum.

It is anticipated that better results could be obtained with this new nonstationary analyzer as compared with other analyzers. The fact that other analyzers do not include the filtering operation represented by $\ell(t)$ indicates that maximum noise smoothing is not being accomplished. Moreover, other analyzers are not generally optimized to take maximum advantage of the limited amount of raw data available.

Experiments have been performed which show that good correspondence exists between the mathematical aspects of the theory and their counterparts in practice. Moreover, for a typical example, but one for which the analyzer was not optimized, the analyzer continued to operate acceptably.

RECOMMENDATIONS FOR FURTHER WORK

Throughout the course of performing the research and experimentation of a project, the investigator invariably "uncovers" new work areas and more work in the same areas. Typically this new or extension type work is beyond the scope of the available time and allotted funds for contract. The following areas of work are recommended as extensions to the research summarized in this report.

1. Nonstationary Processes
   a) Expand the applicability of the nonstationary correlator
   b) Treat the NASA data with the nonstationary correlator

2. Process Testing
   a) Test for stationarity
   b) Test for class of process, e. g. Bivariate gaussian, Rayleigh, etc.
   c) Test of accuracy of estimates

3. Flow Chart of Generalized Processor
   a) Decision making logic
   b) Accuracy realized by each processor
   c) Time saving realized

4. Application of Special Purpose Computers
   a) Use of analog processors for complete data analysis
   b) Use of analog equipment for preprocessing of data
   c) Use of hybrid equipment

5. Digital Correlation Function Computation
   a) Stielje's correlator. Rather than use the two levels of the half-polarity correlator, use three, five or more levels.
   b) Design high speed computer programs to compute Stielje's correlation.
   c) Further study of the half-polarity correlator. Investigate applicability to processes other than bivariate gaussian.
   d) Further study and hypothesis testing of the addition of white noise prior to the half-polarity correlator.

6. Transfer Function Estimation
   a) Determine the best m^{th} order estimate of a system transfer function.
   b) Apply to rocket vibration data to determine an analytic model of the system. This information vitally needed for booster control design also this information should be compared with models obtained from design considerations.

APPENDIX A
OPTIMUM FILTER WEIGHTS

Consider the following ideal frequency function

$$F_i(\omega) = \sum_{\ell=0}^{L} K_\ell \, F_{i\ell}(\omega) \tag{A-1}$$

where

$$F_{i\ell}(\omega) = \sum_{n=-\infty}^{\infty} f_{\ell n} \, e^{-jn\omega T} \tag{A-2}$$

It is desired to obtain an approximate frequency function which has a finite number of terms in its weighting sequence say $2N+1$ terms. The actual realizable filter will be specified by minimizing the following error index

$$J = \int_{-\omega_c}^{\omega_c} \left| F_a(\omega) - F_i(\omega) \right|^2 d\omega \tag{A-3}$$

where the realizable frequency function is given by

$$F_a(\omega) = \sum_{n=-N}^{N} C_n \, e^{-jn\omega T} \tag{A-4}$$

The error index can also be expressed in the following form

$$J = \int_{-\omega_c}^{\omega_c} E(\omega) \, E^*(\omega) \, d\omega \tag{A-5}$$

where $E(\omega) = F_a(\omega) - F_i(\omega)$

now let $z = e^{j\omega T}$

and the error index becomes

$$J = \frac{\omega_s}{2\pi j} \oint_\Gamma E(z) \, E(z^{-1}) \, z^{-1} \, dz \tag{A-6}$$

where the contour Γ is the unit circle. Now using the discrete form of Parseval's theorem we have

$$J = \omega_s \sum_{n=-\infty}^{\infty} e_n^2$$

$$= \omega_s \sum_{n=-\infty}^{\infty} \left(C_n - \sum_{\ell=0}^{L} K_\ell f_{\ell n} \right)^2 \qquad \text{(A-7)}$$

where $C_n = 0 \; , \; |n| > N$

Minimizing with respect to the C's yield the following

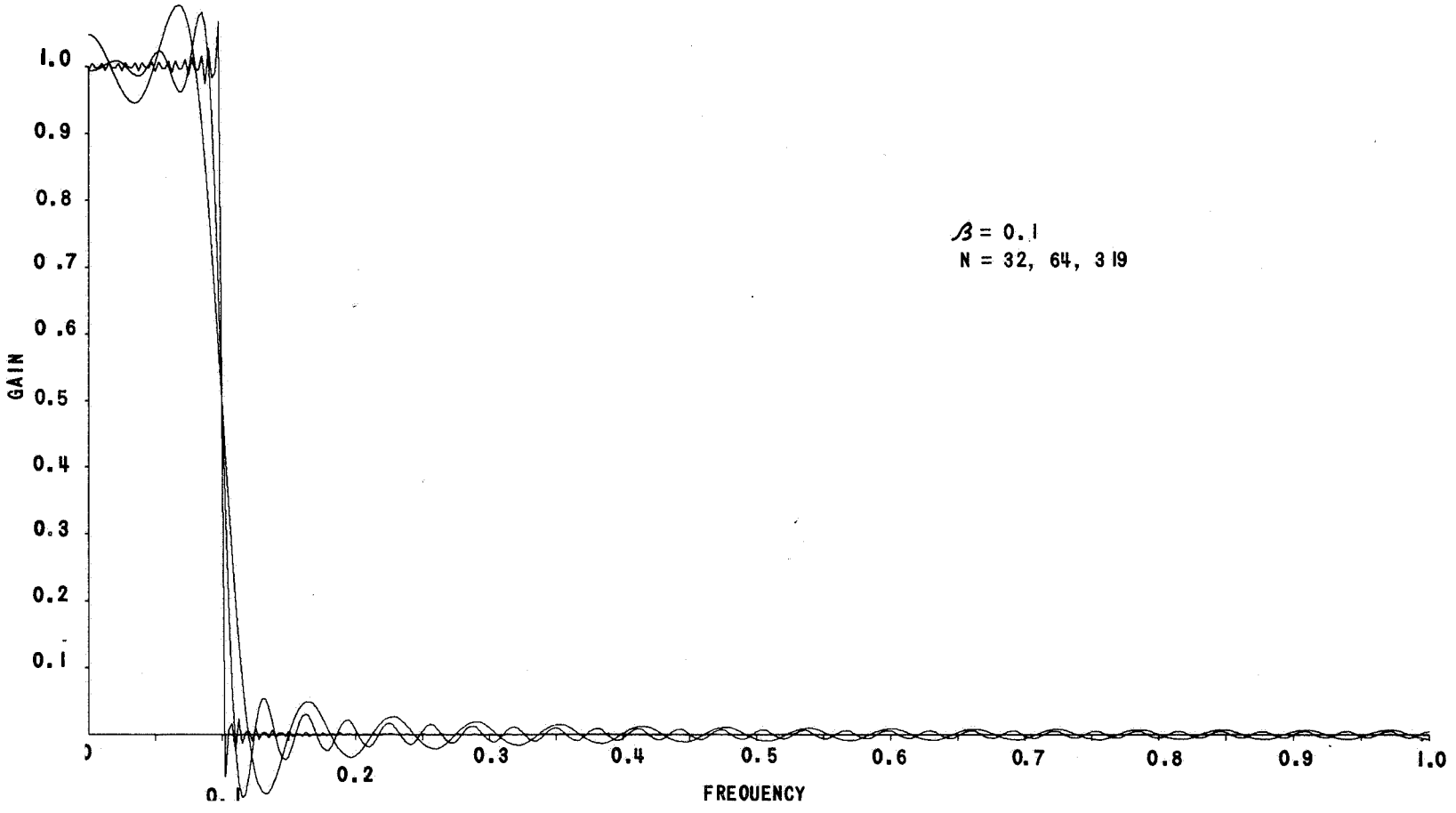
$$C_n = \sum_{\ell=0}^{L} K_\ell f_{in} \qquad |n| \leq N \qquad \text{(A-8)}$$

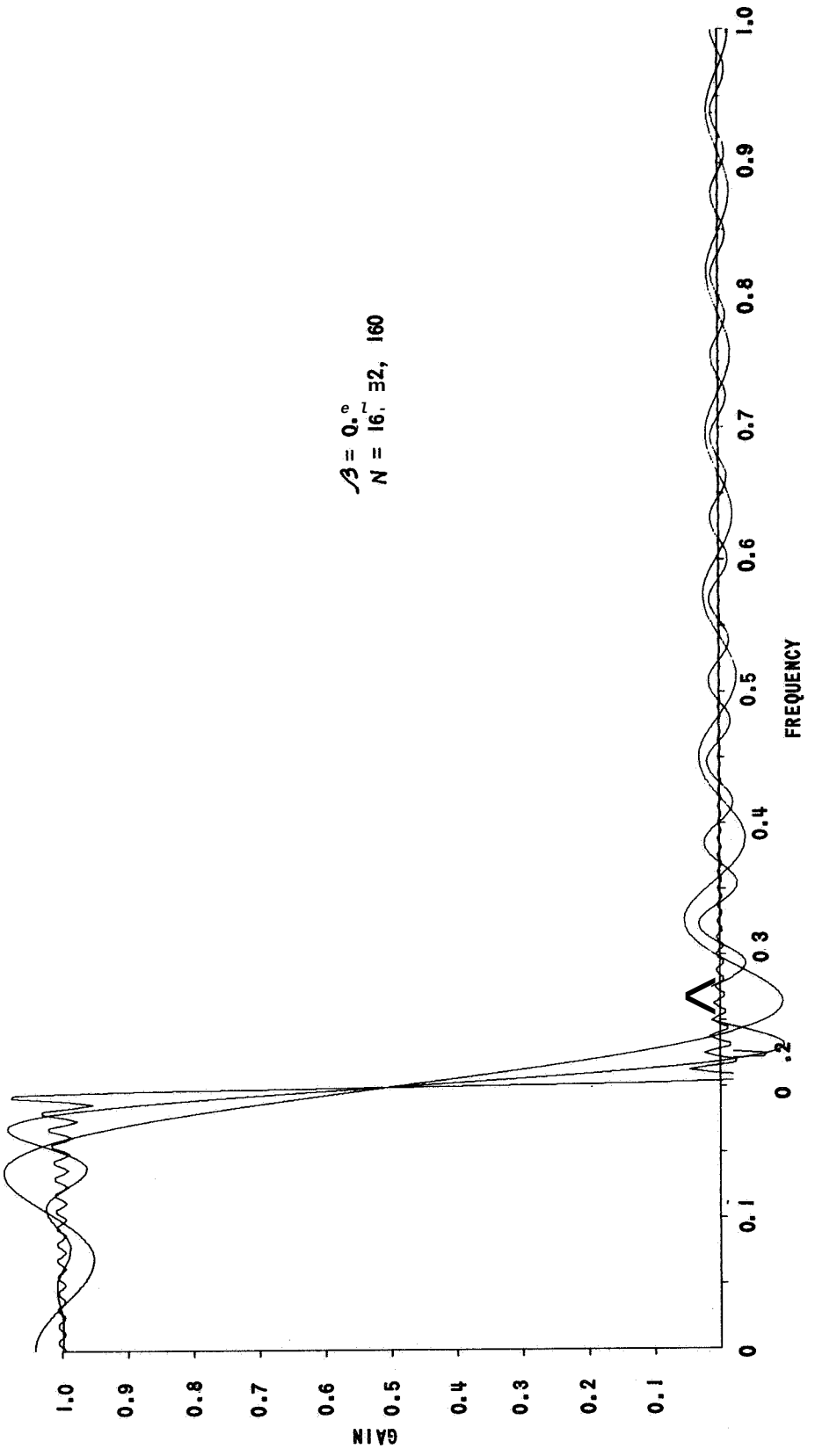
Equation (A-8) indicates that the optimum realizable filter is obtained by taking the ideal weights intact that is no modification is needed due to the fact that only a finite number of weights are used. Another interesting aspect of equation (A-8) is that it indicates that the same number of weights should be used by each component filter.

APPENDIX B
LOW PASS FREQUENCY RESPONSES

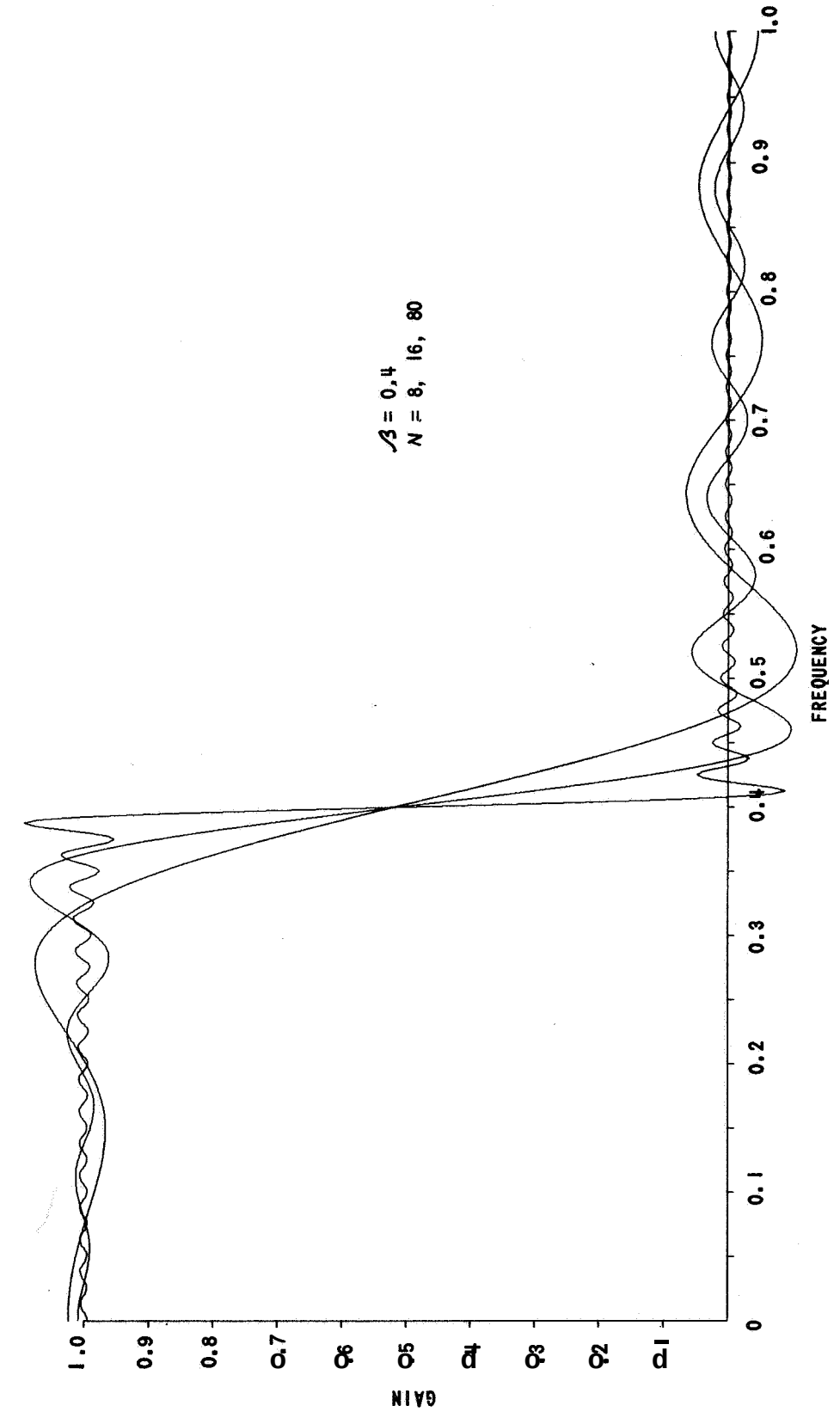
The relationship given below permits the frequency plots on the following pages to be inverted to visually produce a second set of 27 graphs.

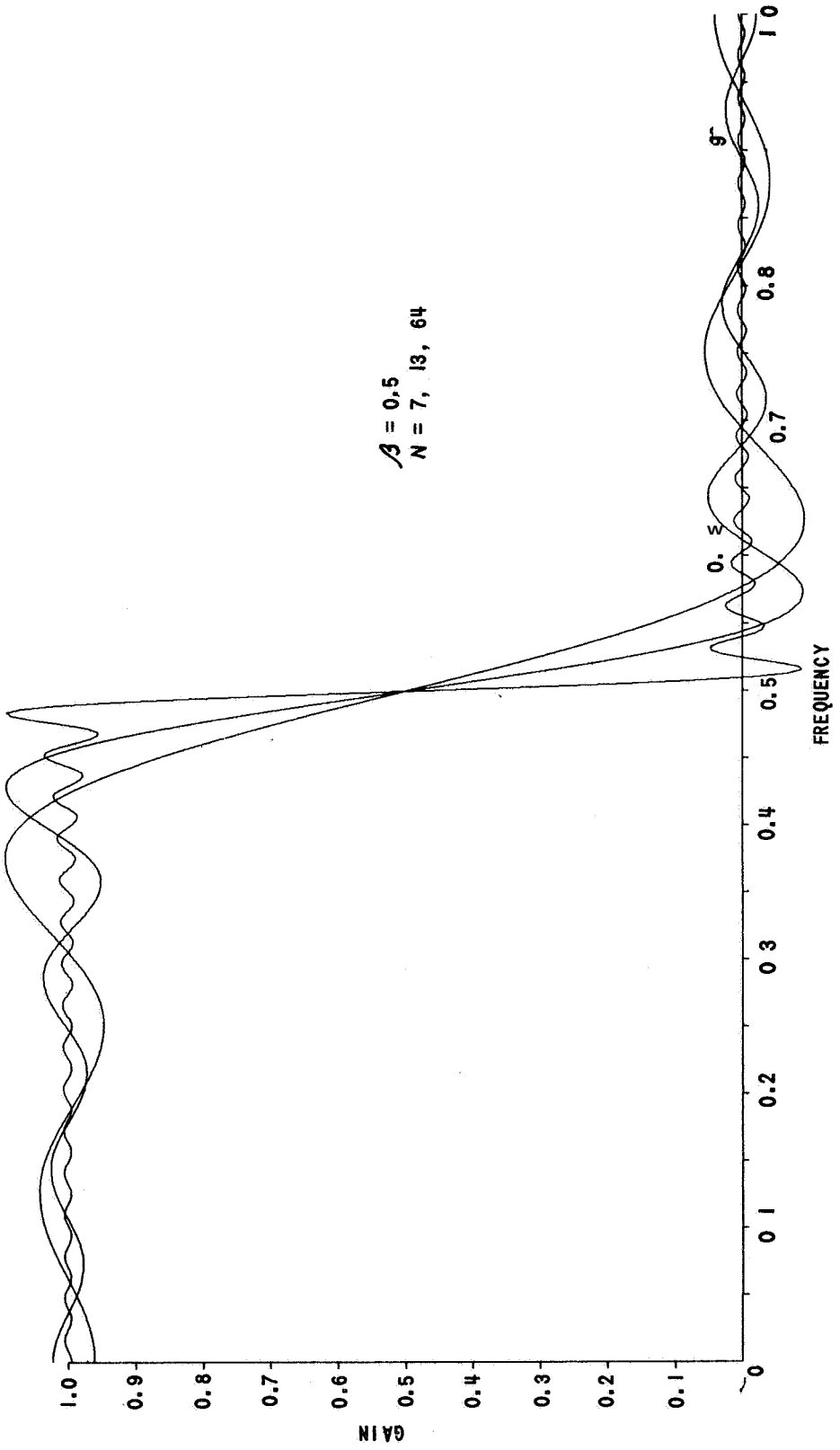
$$F(\alpha, \beta, N) = 1 - F(1+\alpha, 1-\beta, N)$$

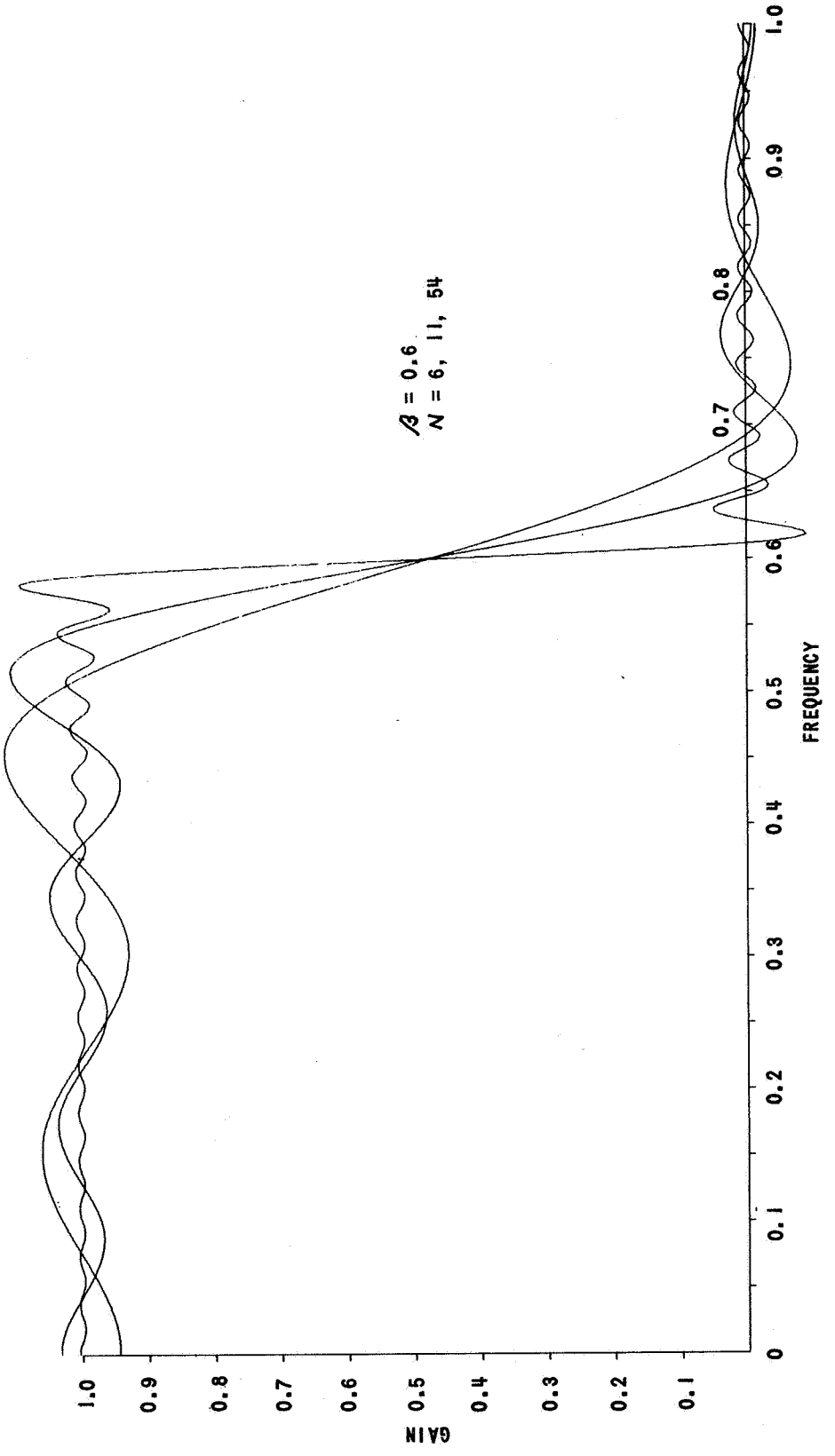

1.0
0.9
0.8
0 .7
0 .6
0.5
0.4
0.3
0.2
0.1
GAIN
0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
FREQUENCY
β = 0.1
N = 32, 64, 319

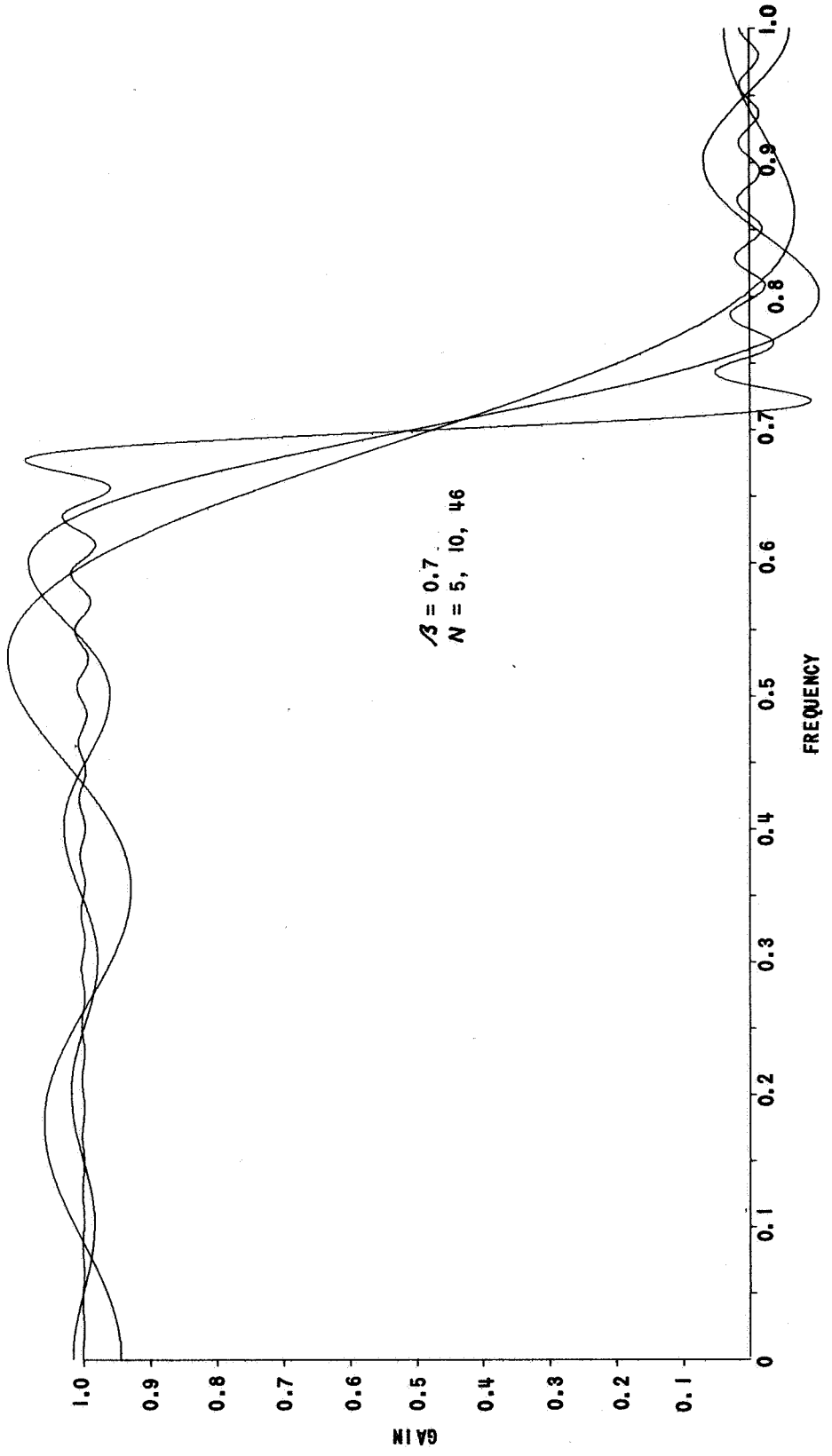

β = 0.
N = 16, 32, 160
GAIN
FREQUENCY
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
.2

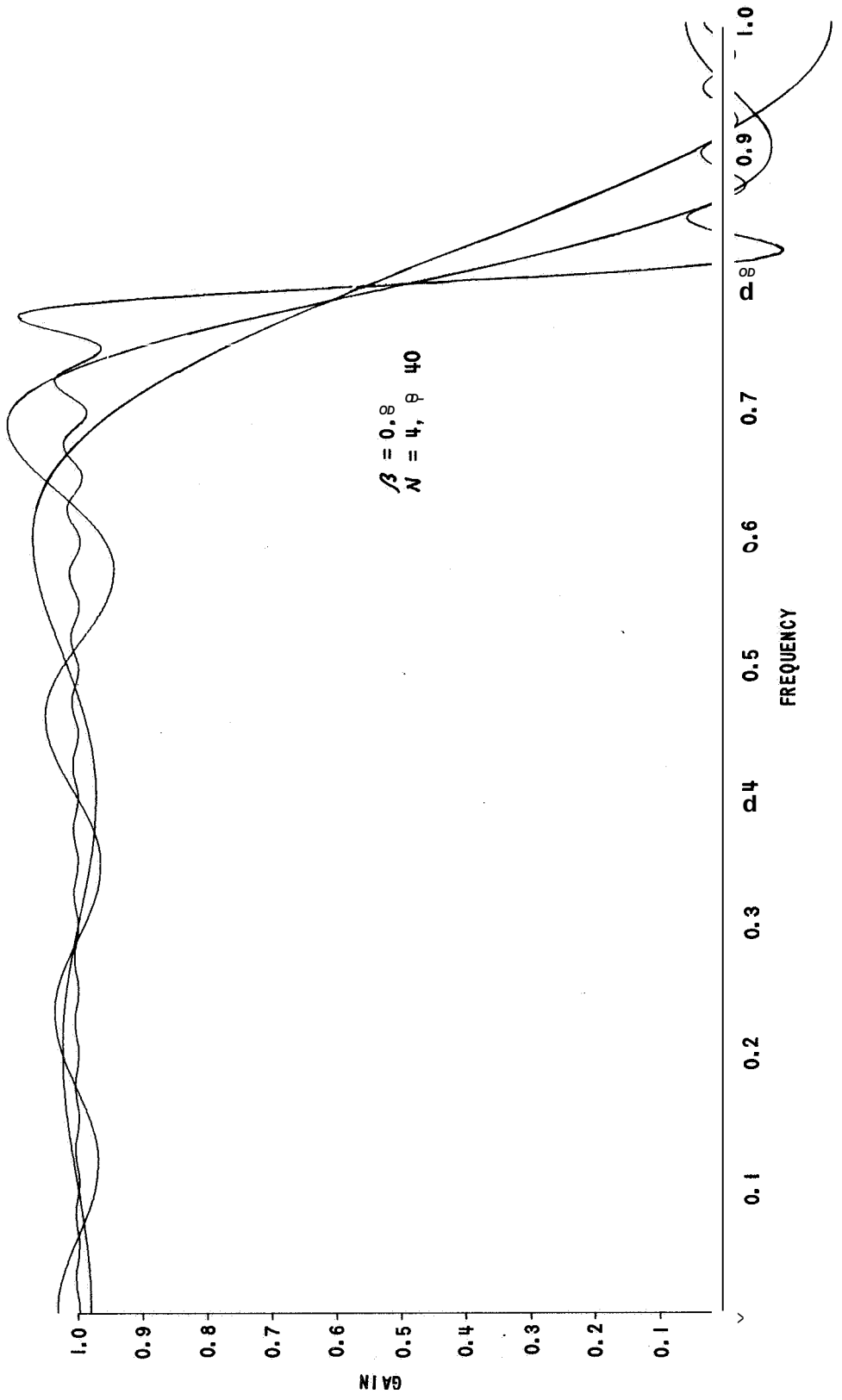
β = 0.3
N = 11, 22, 107
GAIN
FREQUENCY
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

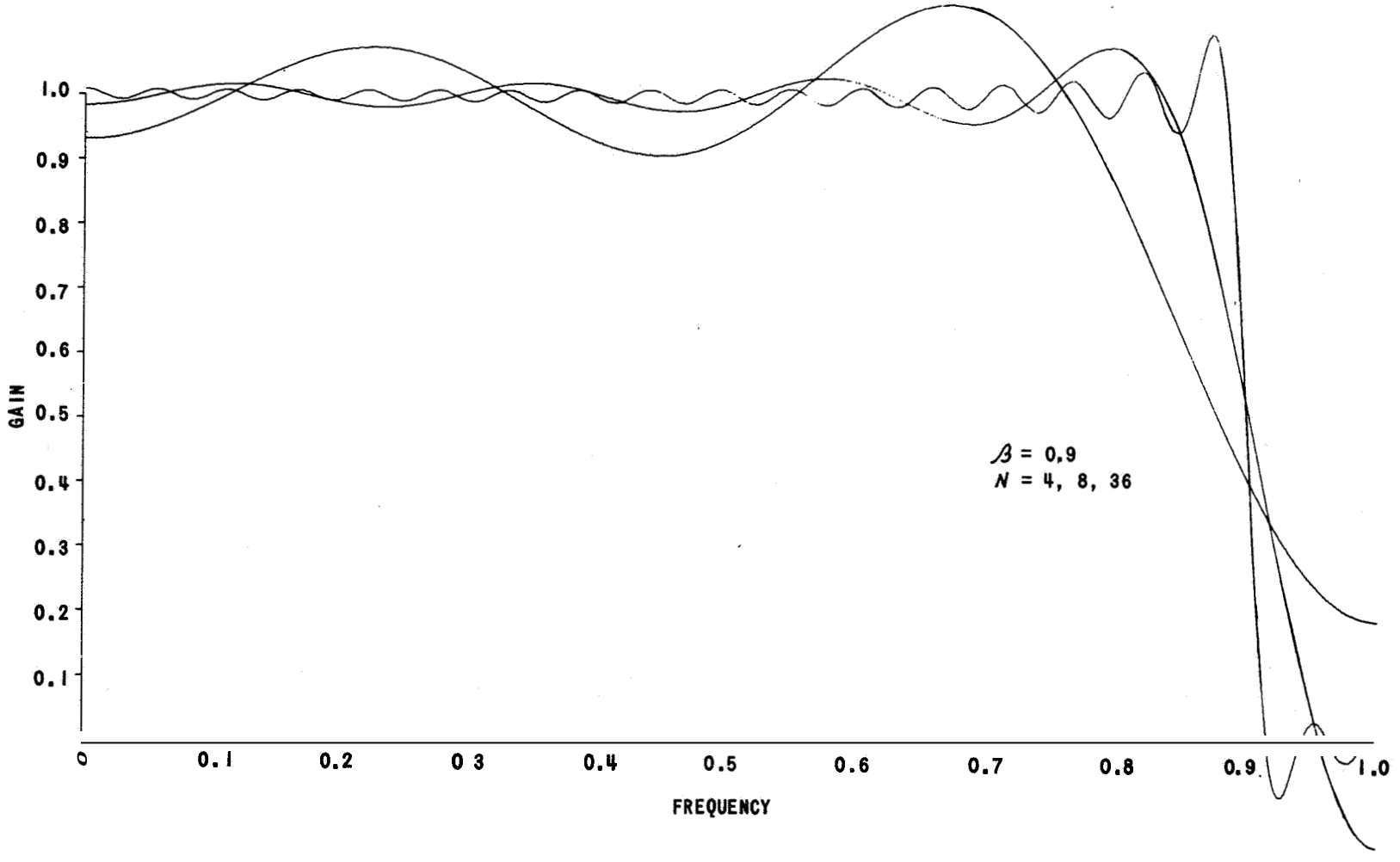

β = 0.4
N = 8, 16, 80
GAIN
FREQUENCY
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


β = 0,5
N = 7, 13, 64
GAIN
FREQUENCY
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.2
0.3
0.4
0.5
0.7
0.8
1.0


β = 0.6
N = 6, 11, 54
GAIN
FREQUENCY
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


GAIN
FREQUENCY
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
β = 0.7
N = 5, 10, 46


GAIN
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.9
1.0
FREQUENCY
dB
β = 0.08
N = 4, 8, 40


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
GAIN
β = 0,9
N = 4, 8, 36
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
FREQUENCY

APPENDIX C
OPTIMUM FILTER WEIGHTS WITH UNITY **D.C.** GAIN

With the D. C. gain constraint added, the error index becomes

$$J = \int_{-\omega_c}^{\omega_c} \left| F_a(\omega) - F_c(\omega) \right|^2 d\omega + \lambda \left(\sum_{n=-N}^{N} C_n - 1 \right)$$

$$\sum_{n=-\infty} e_n \qquad \lambda \left(\sum_{n=-N} C_n - 1 \right)$$

$$\sum_{n=-\infty}^{\infty} \left(C_n - \sum_{\ell=0}^{L} K_\ell f_\ell \right) + \lambda \left(\sum_{n=-N} C_n - 1 \right) \qquad \text{(C-1)}$$

Minimizing the above index with respect to C_i yields

$$\frac{\partial J}{\partial C_i} = 2\omega_s \left(C_i - \sum_{\ell=0}^{L} K_\ell f_{\ell i} \right) + \lambda \; ; -N \le i \le N \qquad \text{(C-2)}$$

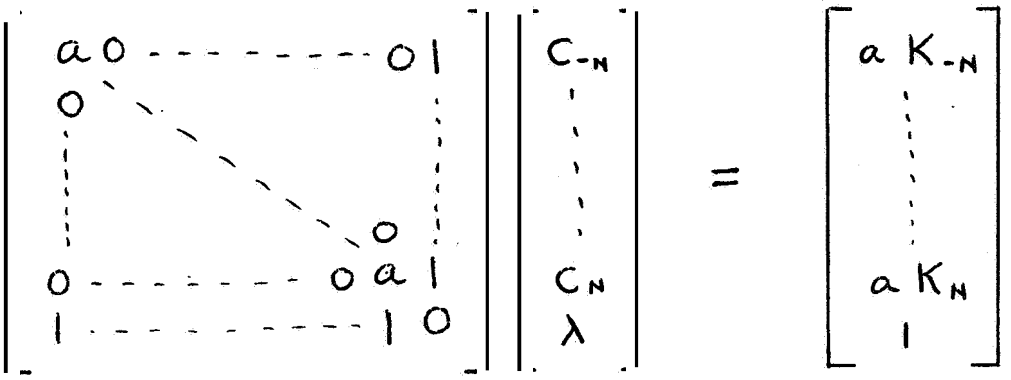
The above $2N+1$ equations a ong with the constraint equation comprise a set of $2N+2$ linearly independent equations in $2N+2$ unknowns whose solution is obtained as follows

let

$$a = 2\omega_s$$

$$K_n = \sum_{\ell=0}^{L} K_\ell f_{\ell n}$$

Placing the $2N+2$ equations into matrix form yields



Partitioning the above matrix and expanding yields

$$Ac + b\lambda = k \tag{C-3}$$

$$b'c = 1 \tag{C-4}$$

where

$$A = \begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$k = \begin{bmatrix} aK_{-N} \\ \\ aK_N \end{bmatrix} \qquad c = \begin{bmatrix} C_{-N} \\ \vdots \\ C_N \end{bmatrix} \tag{C-5}$$

Utilizing the above notation, the solution for the optimum weights are obtained from (C-3) and (C-4) as follows:

$$c + A^{-1} b \lambda = A^{-1} k$$

$$b' c + b' A^{-1} b \lambda = b' A^{-1} K$$

$$1 + b' A^{-1} b \lambda = b' A^{-1} K$$

$$\lambda = (b' A^{-1} b)^{-1} (b' A^{-1} K - 1)$$

$$c = A^{-1} K - A^{-1} b (b' A^{-1} b)^{-1} (b' A^{-1} K - 1) \quad \text{(C-6)}$$

From the relations in (C-5) we have

$$A^{-1} = \frac{1}{a} I \quad \text{(C-7)}$$

Substituting (C-7) into (C-6) yields

$$c = \frac{1}{a} K - \frac{1}{a} b \frac{\sum_{n=-N}^{N} K_n - 1}{\frac{1}{a}(2N+1)}$$

$$C_n = K_n - \frac{\sum_{j=-N}^{N} K_j - 1}{2N+1}$$

$$C_n = \sum_{\ell=0}^{L} k_\ell f_{\ell n} + \frac{1 - \sum_{j=-N}^{N} \sum_{\ell=0}^{L} k_\ell f_{\ell j}}{2N+1} \quad \text{(C-8)}$$

APPENDIX D
CORRELATION FUNCTIONS AND POWER SPECTRA

The results of processing the four NASA records -

E127#11
E128#11
E12#3
E12#4

are given in this Appendix. The correlation functions for the full precision and half-polarity correlator are plotted. The respective (Hanned) power spectral densities of the respective normalized correlations are also plotted.

Figure D1 FULL PRECISIO, CORRELATION FUNCTION OF RECORD E127#11

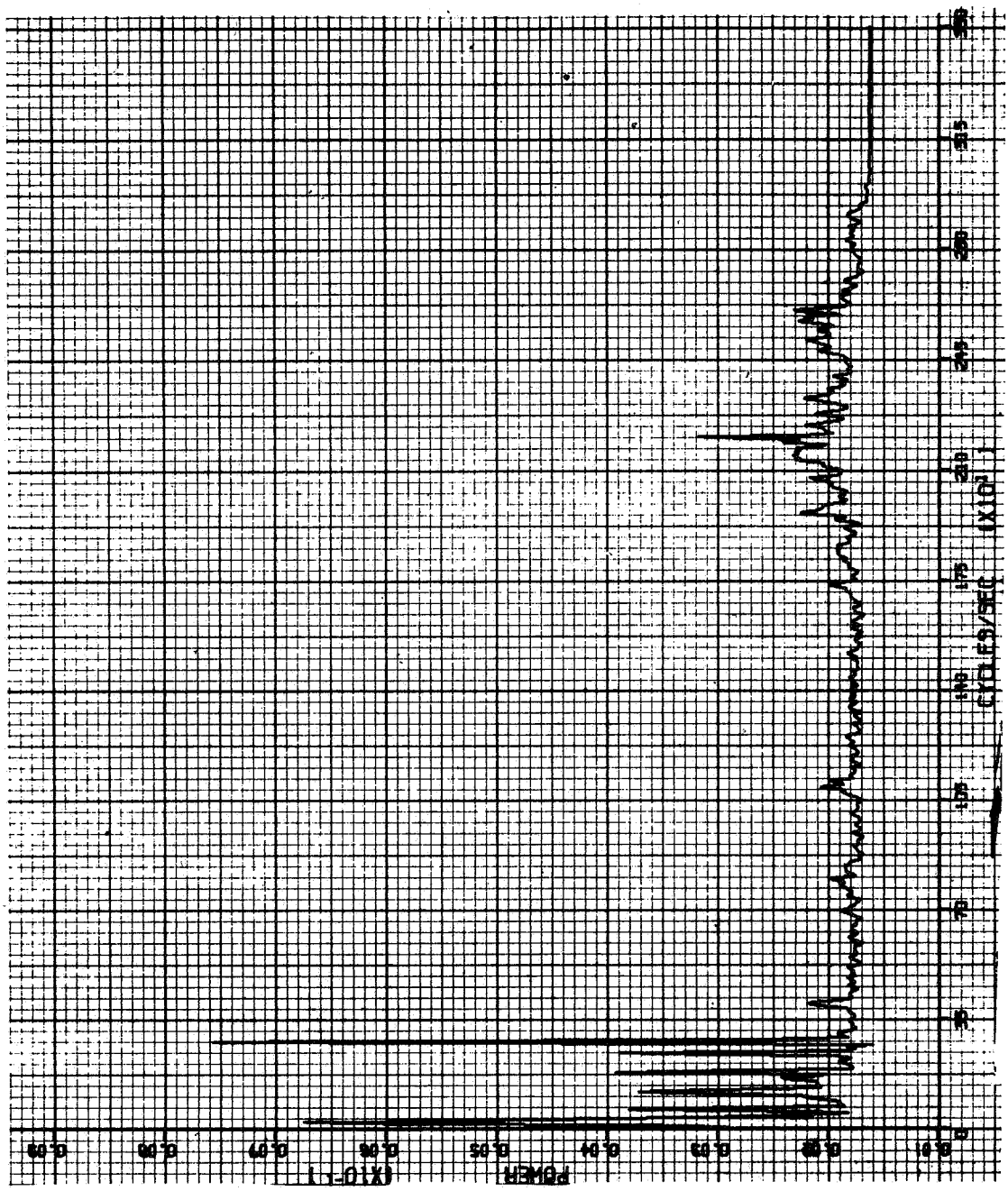


Figure D2 POWER SPECTRUM OF RECORD E127#11 FOR NORMALIZED FULL PRECISION CORRELATOR

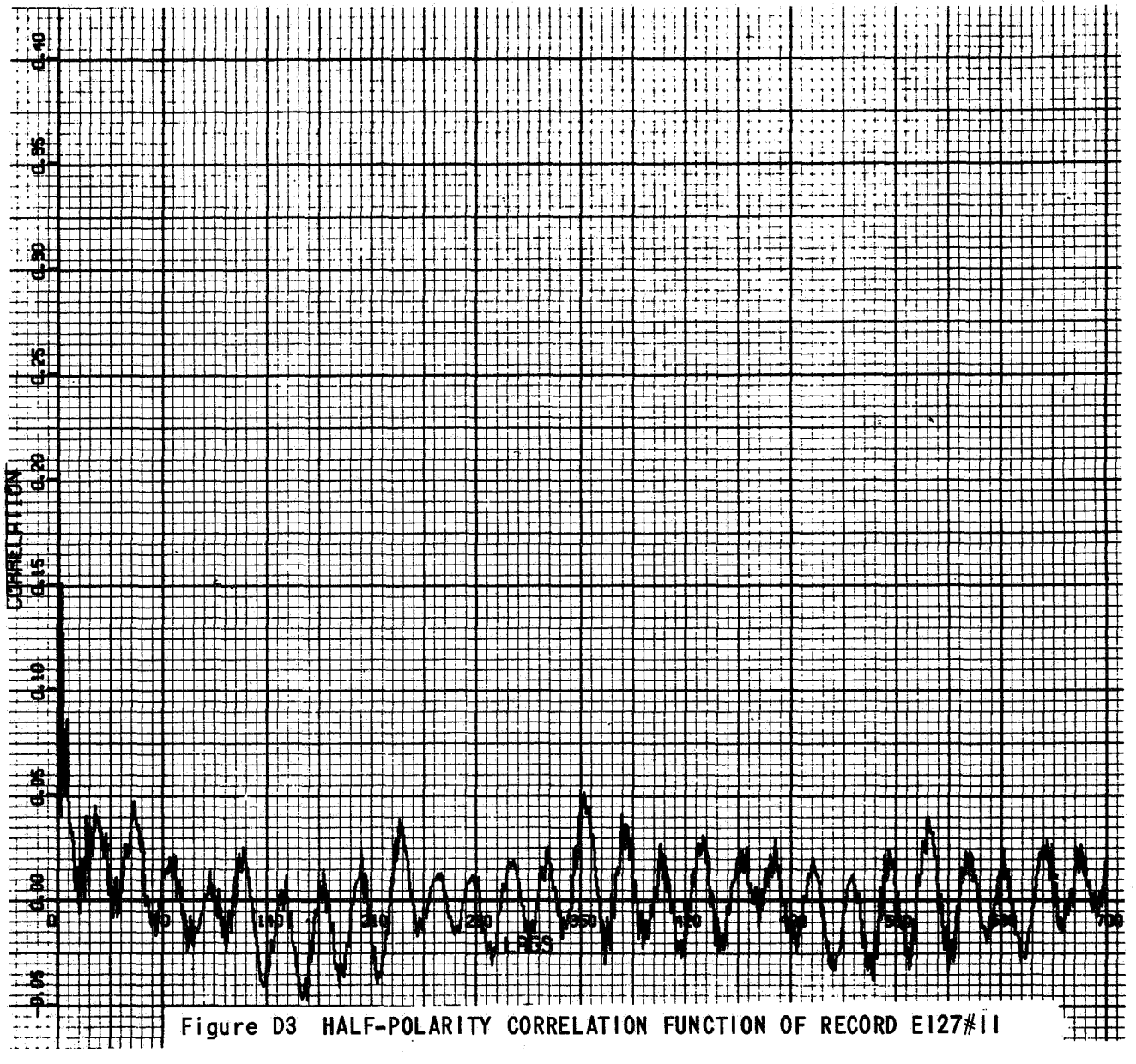


Figure D3 HALF-POLARITY CORRELATION FUNCTION OF RECORD E127#11

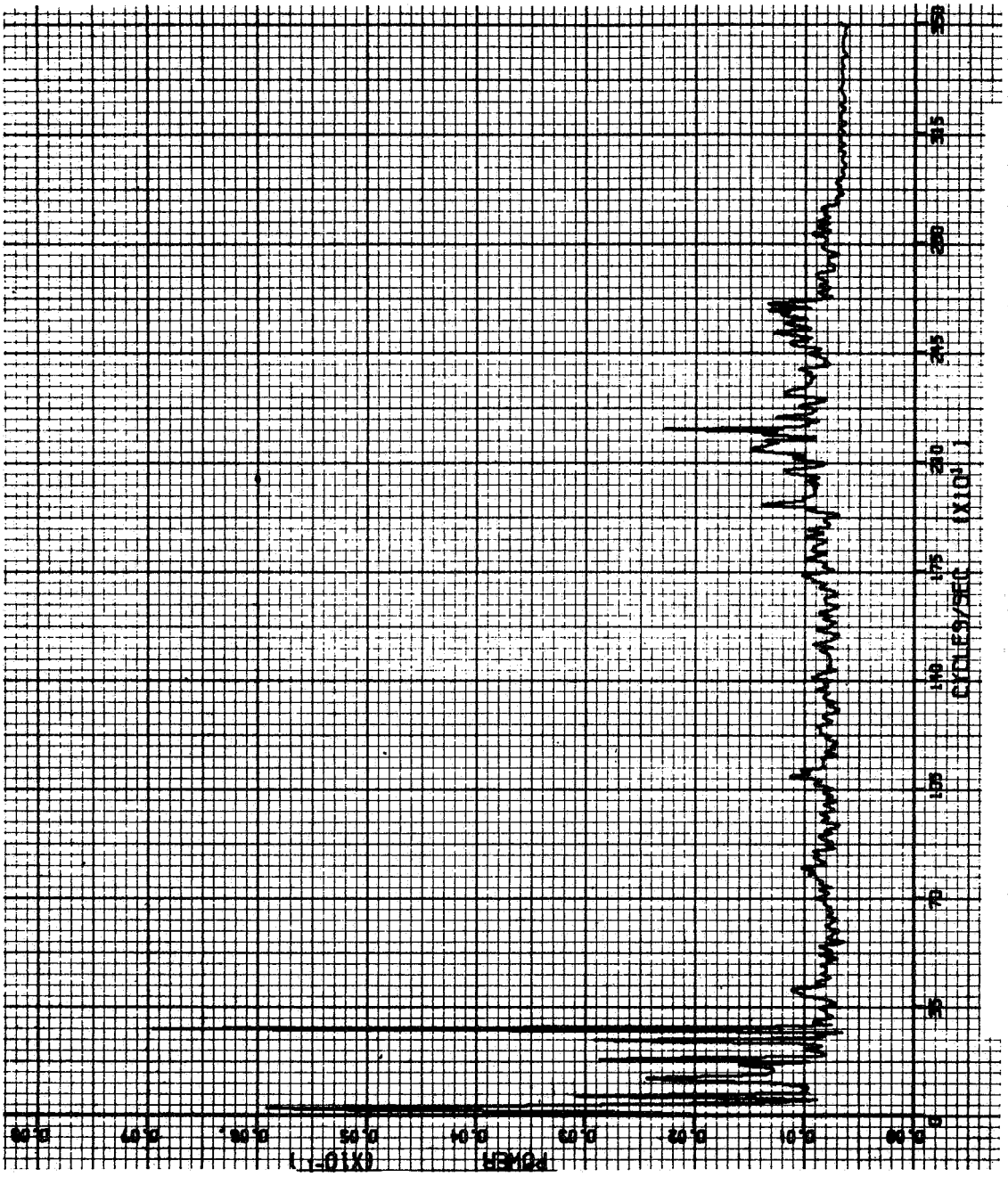


Figure D4 POWER SPECTRUM OF RECORD E127#11 FOR NORMALIZED HALF-POLARITY CORRELATOR

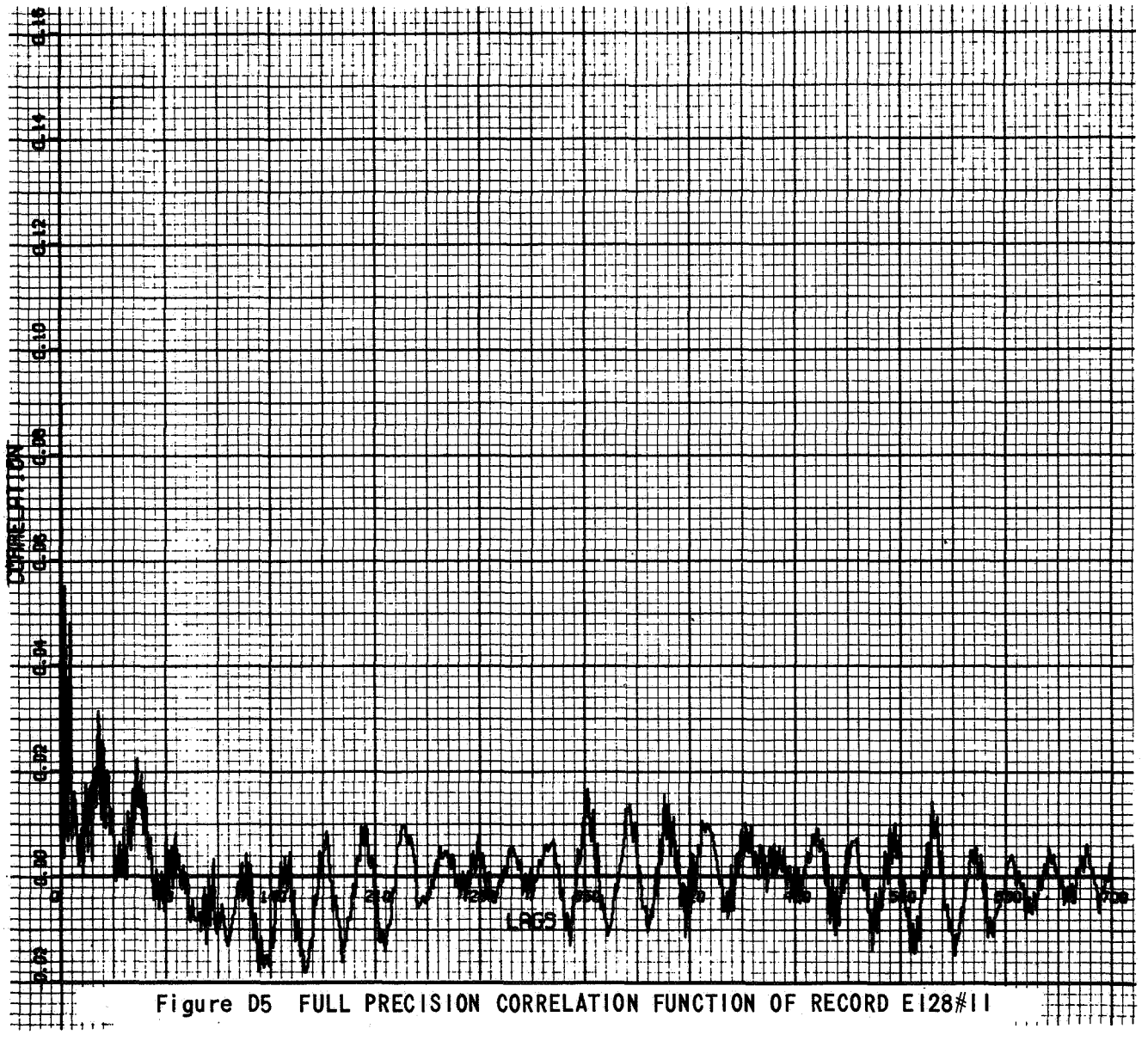


Figure D5 FULL PRECISION CORRELATION FUNCTION OF RECORD E128#11

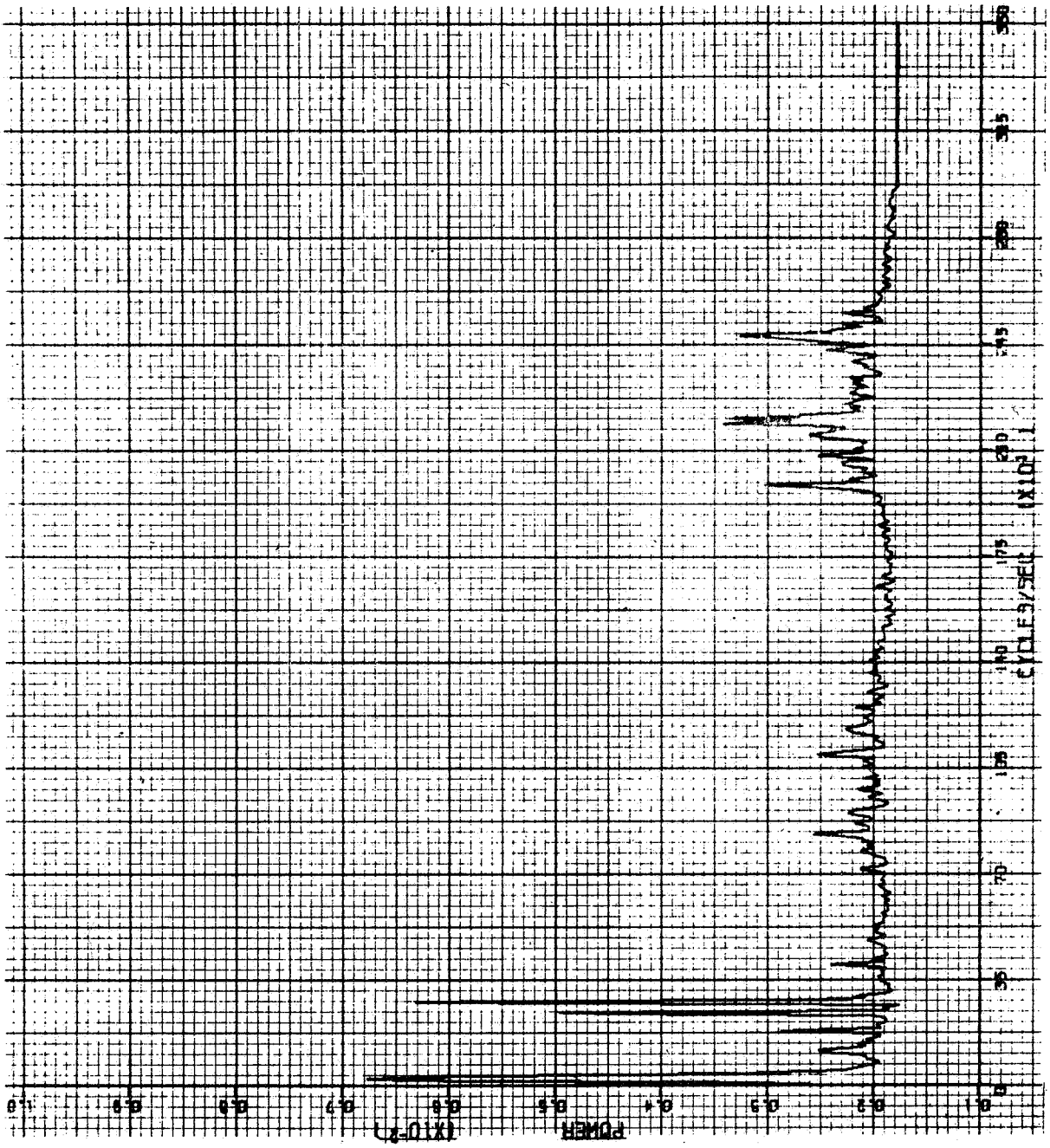


Figure D6 POWER SPECTRUM OF RECORD E128#11 FOR NORMALIZED FULL PRECISION CORRELATOR

Figure D7 HALF-POLARITY CORRELATION FUNCTION OF RECORD E128#11

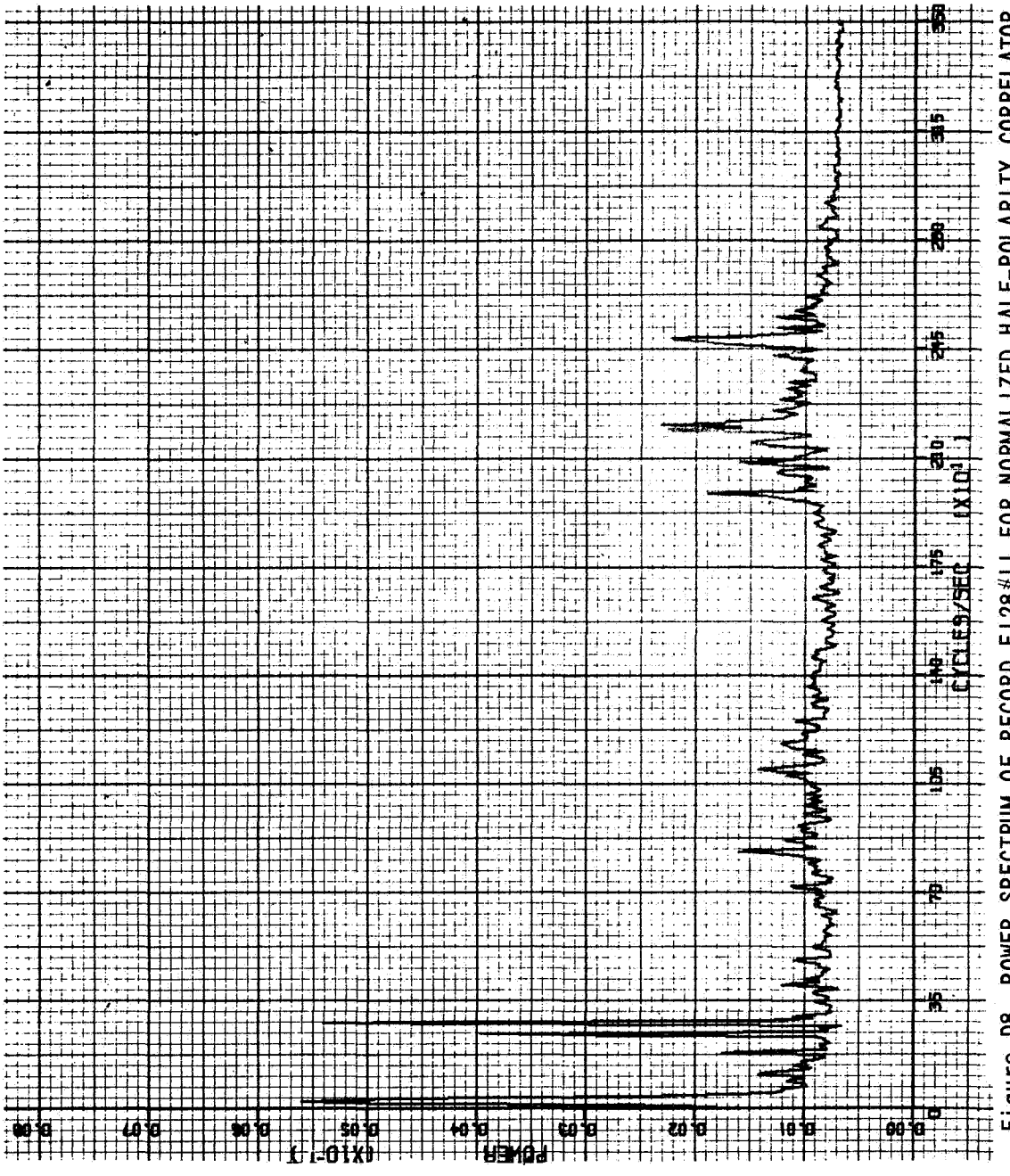


Figure D8 POWER SPECTRUM OF RECORD E128#11 FOR NORMALIZED HALF-POLARITY CORRELATOR

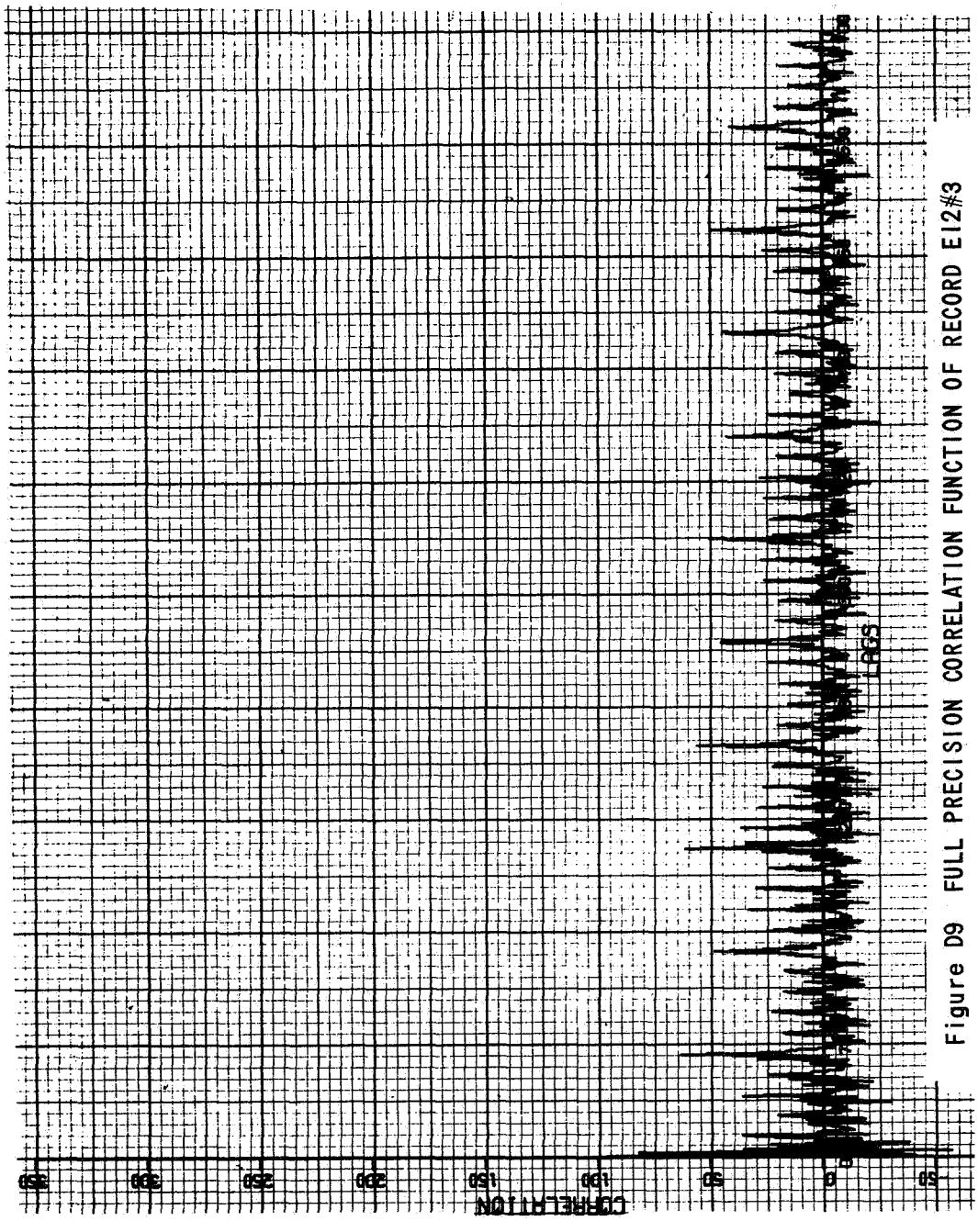


Figure D9 FULL PRECISION CORRELATION FUNCTION OF RECORD E12#3

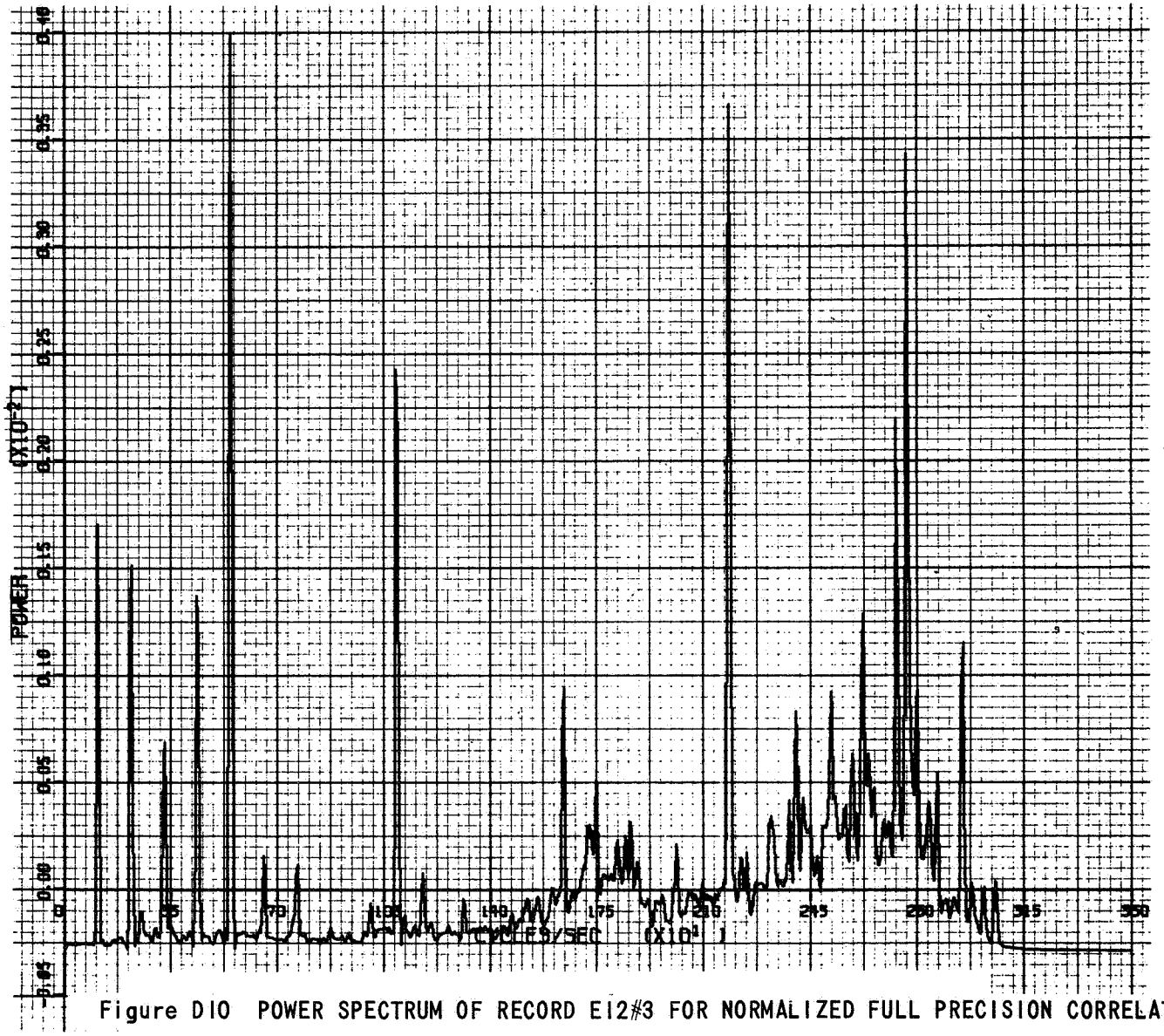


Figure D10 POWER SPECTRUM OF RECORD E12#3 FOR NORMALIZED FULL PRECISION CORRELATOR

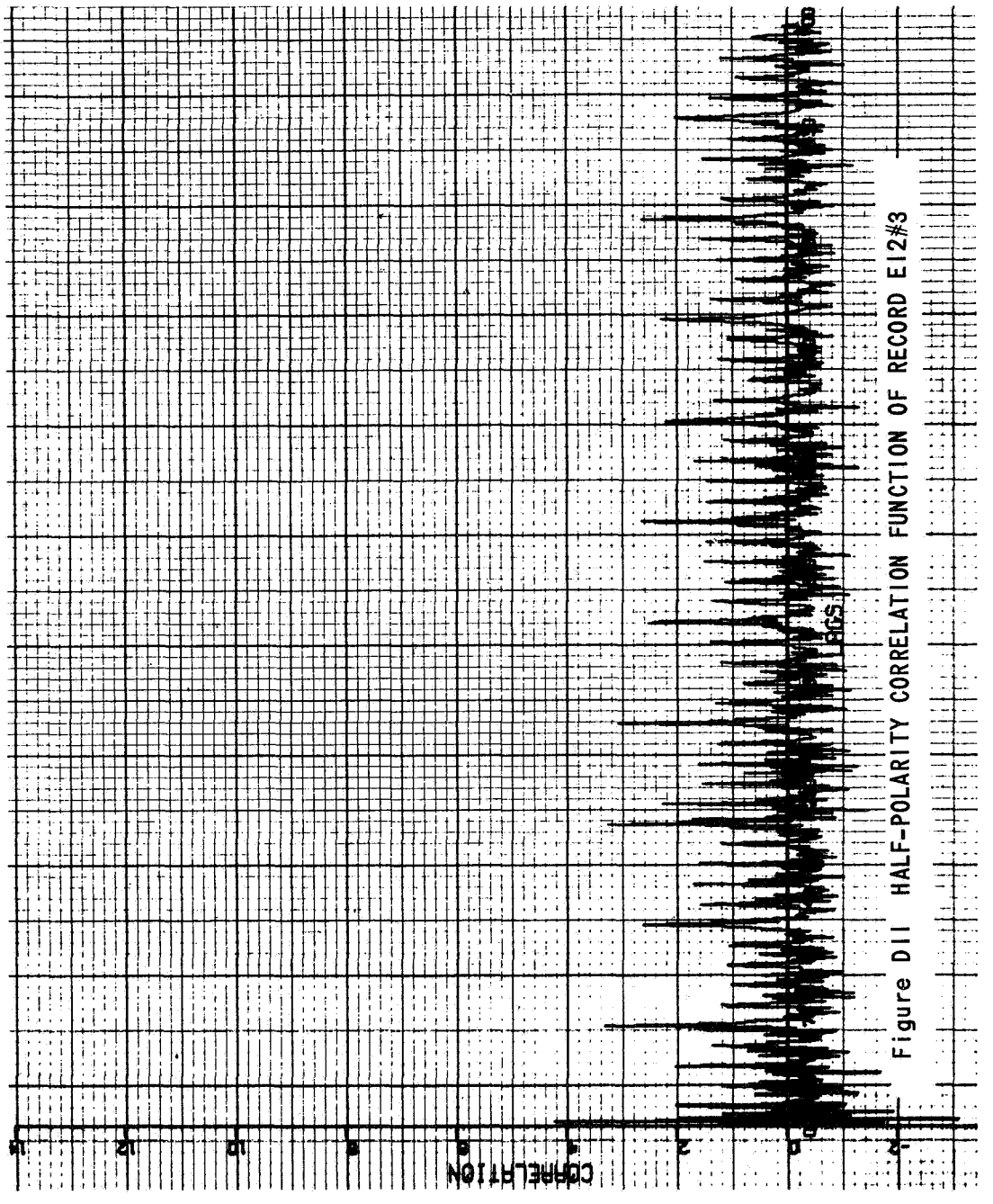


Figure D11 HALF-POLARITY CORRELATION FUNCTION OF RECORD E12#3

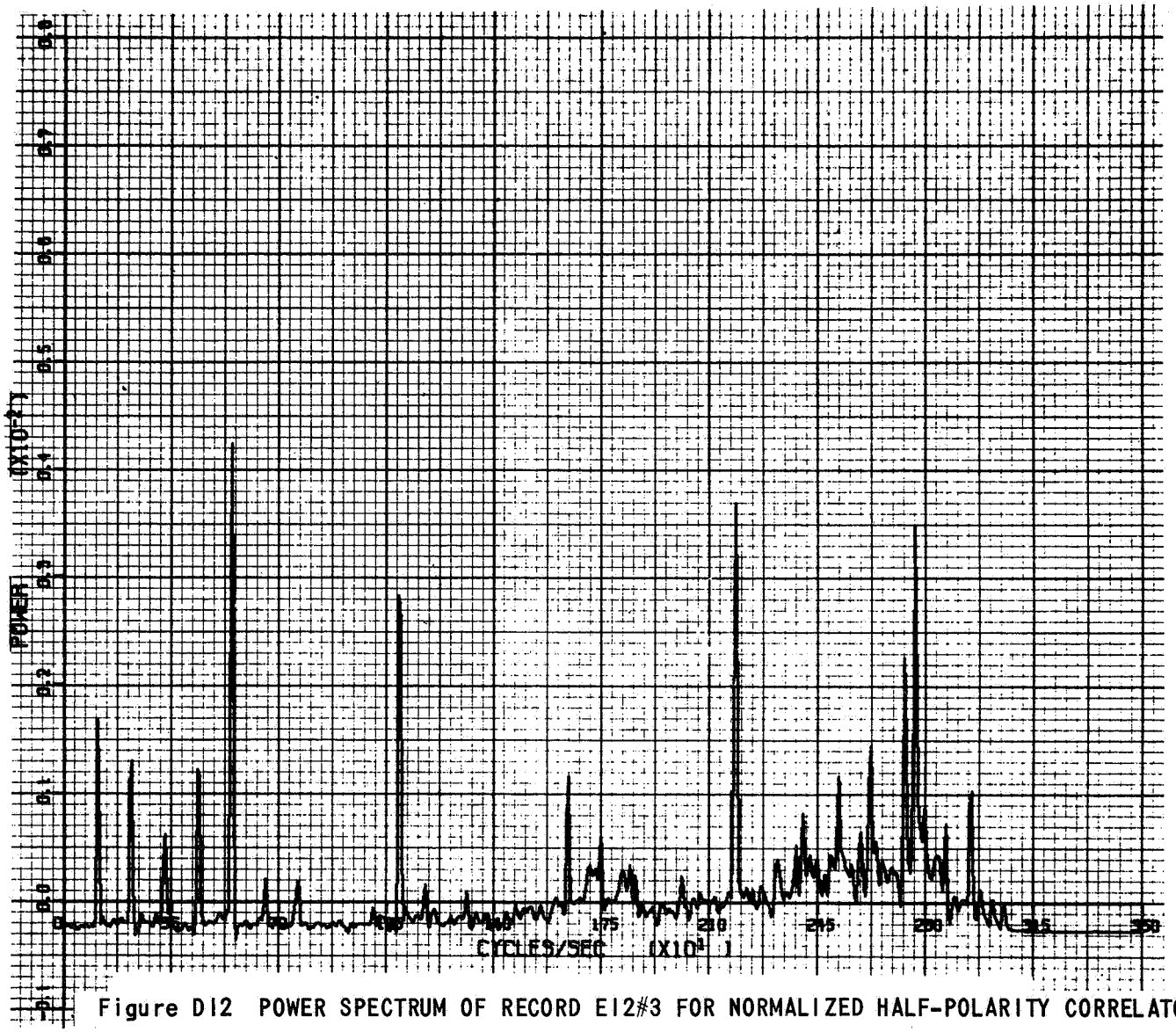


Figure D12 POWER SPECTRUM OF RECORD E12#3 FOR NORMALIZED HALF-POLARITY CORRELATOR

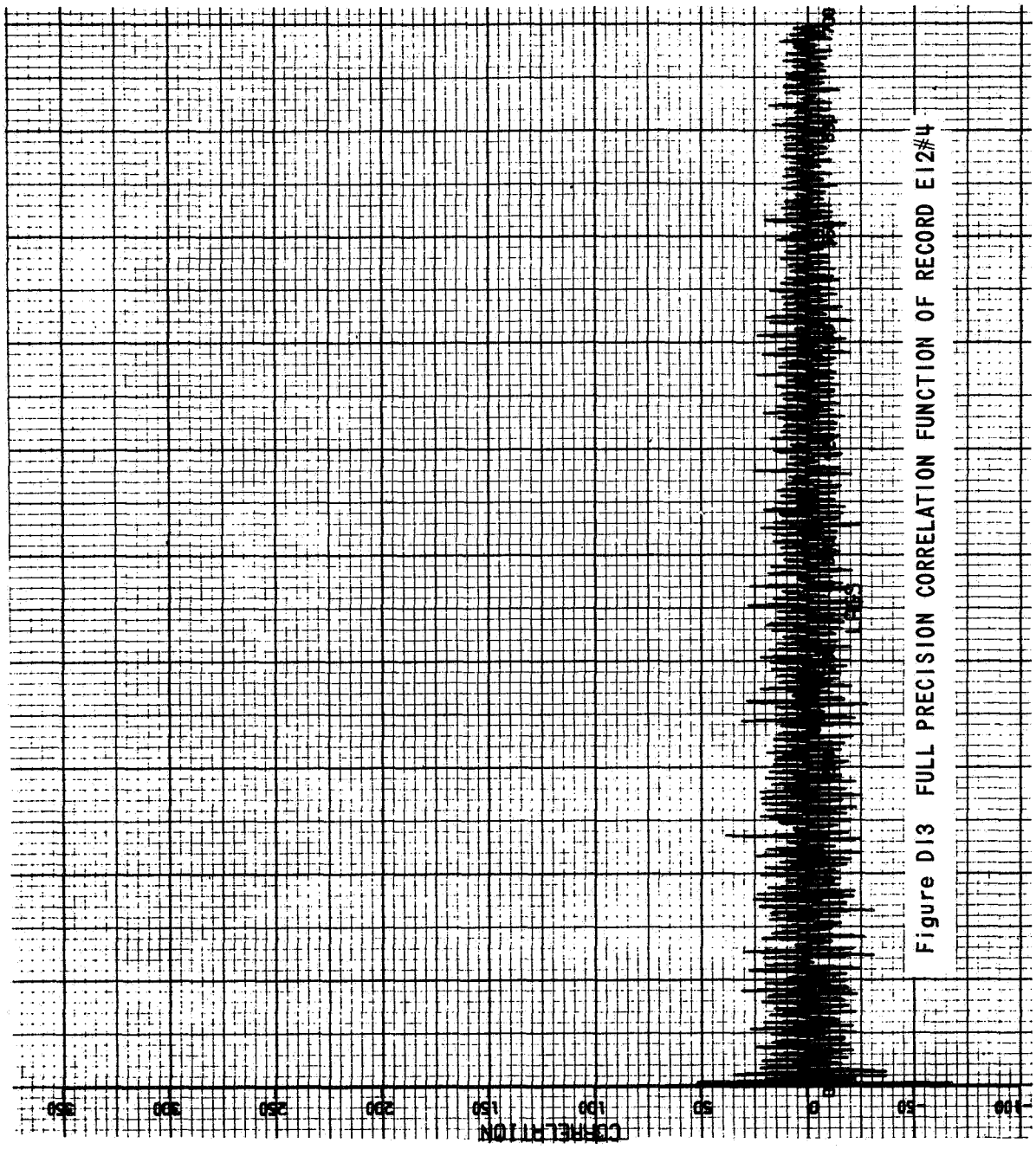


Figure D13 FULL PRECISION CORRELATION FUNCTION OF RECORD E12#4

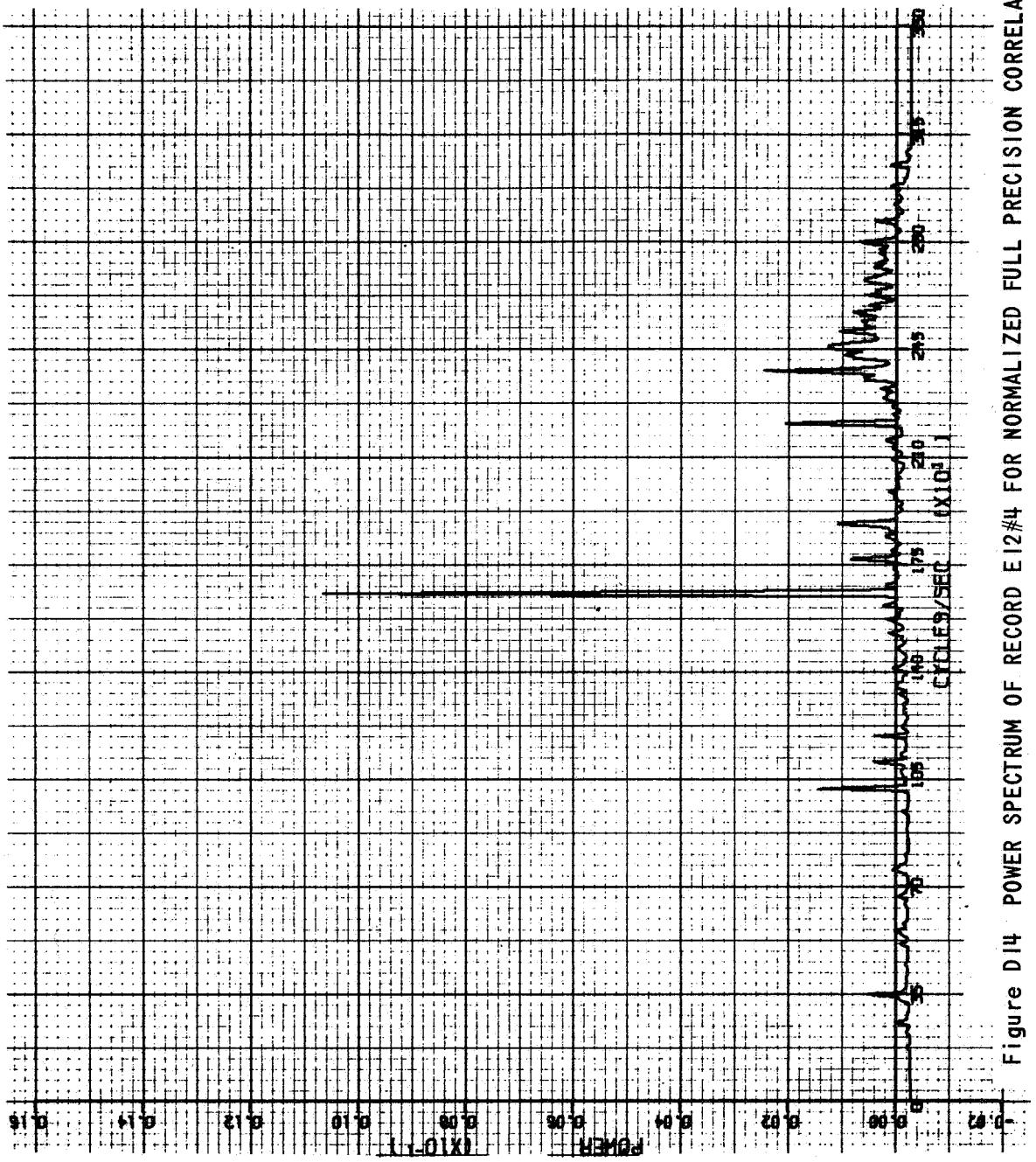


Figure D14 POWER SPECTRUM OF RECORD E12#4 FOR NORMALIZED FULL PRECISION CORRELATOR

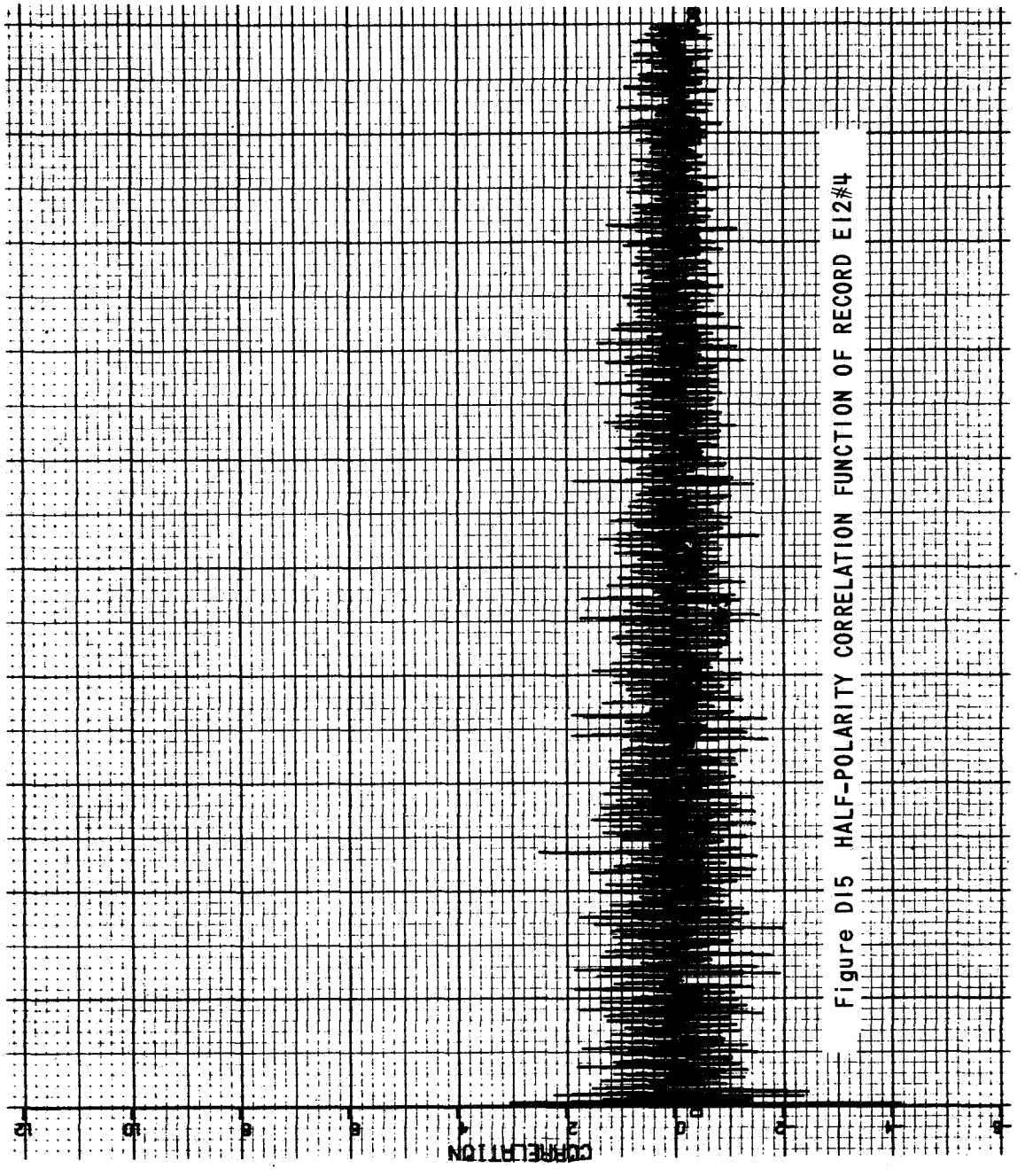


Figure D15 HALF-POLARITY CORRELATION FUNCTION OF RECORD E12#4

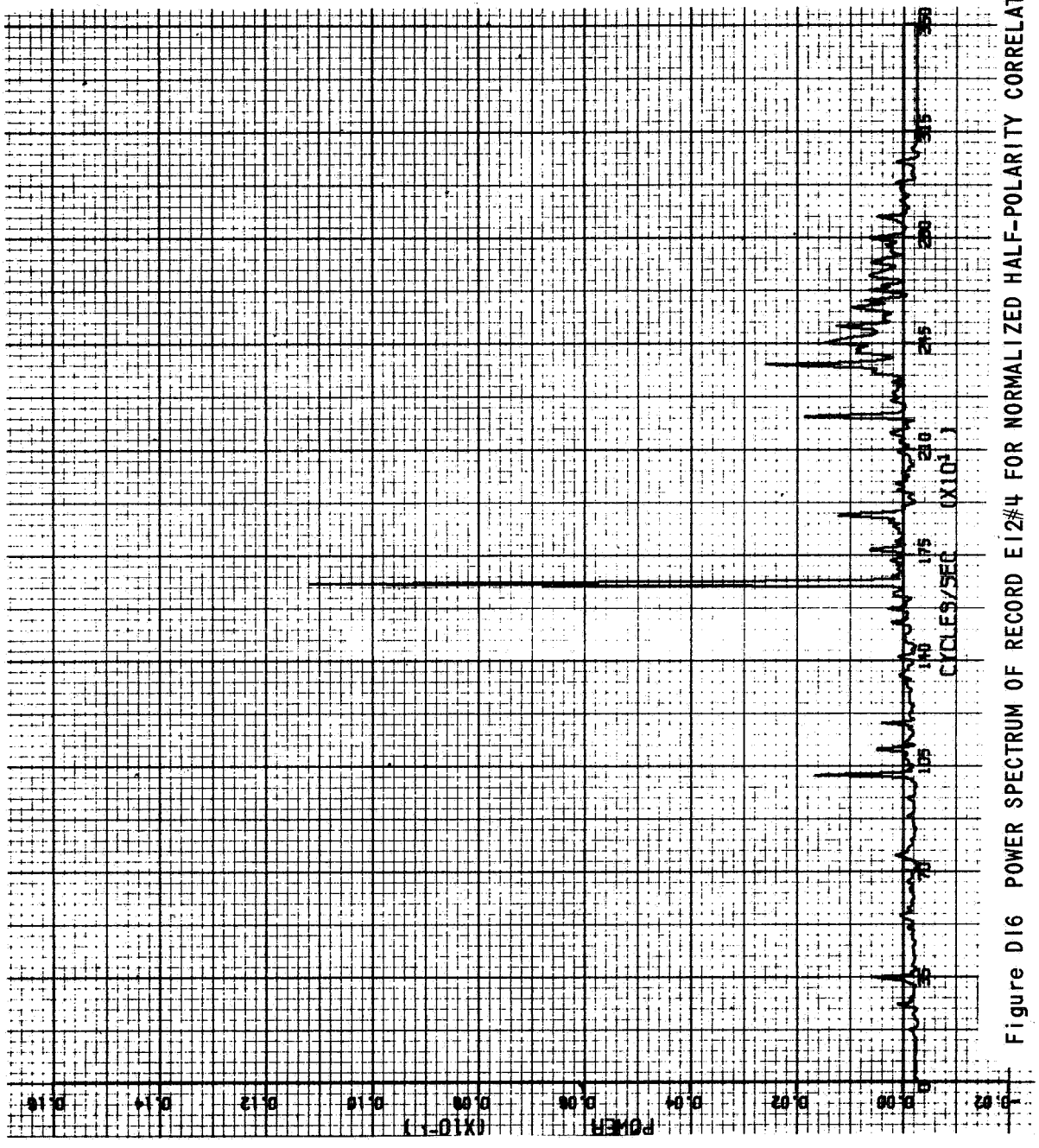
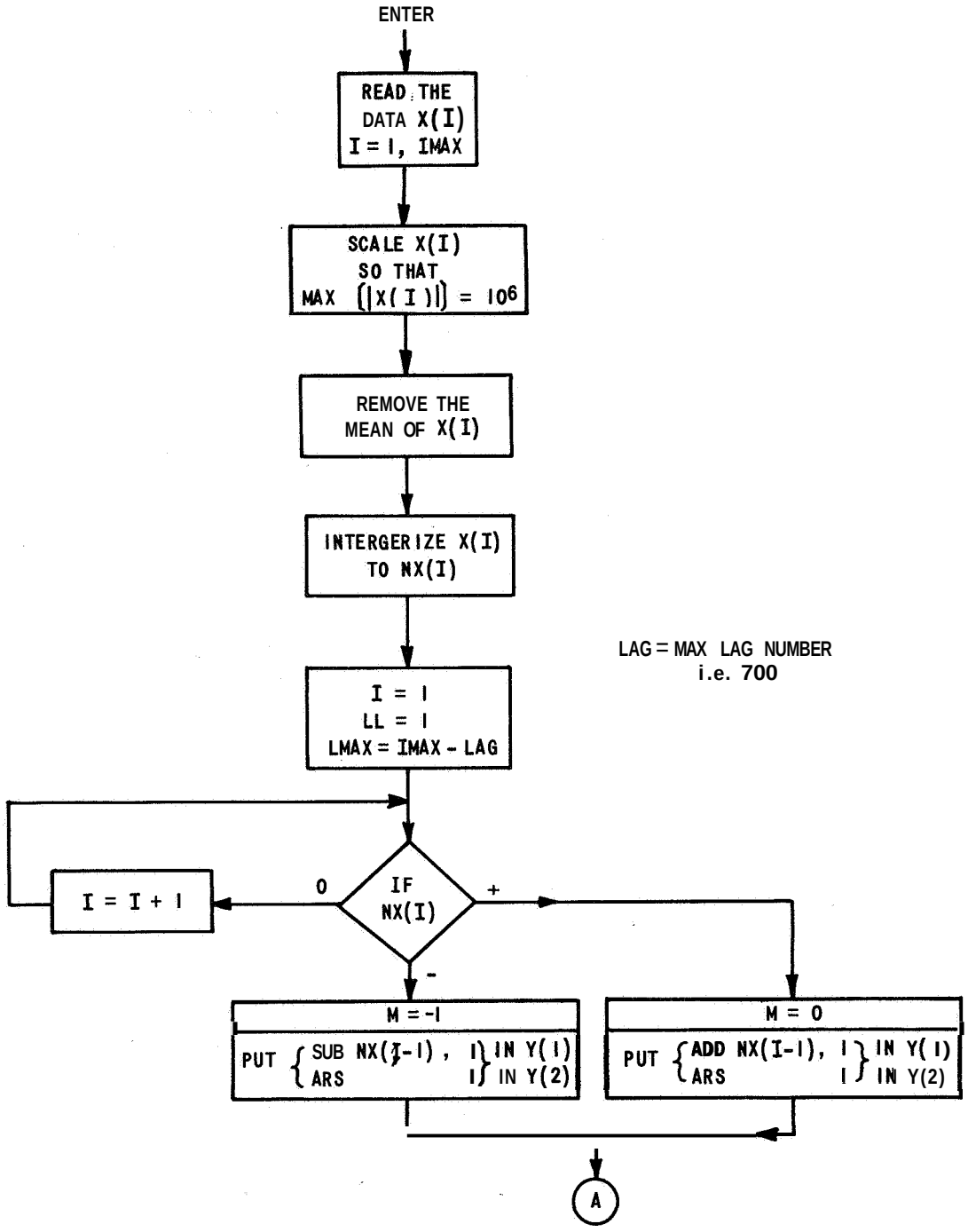
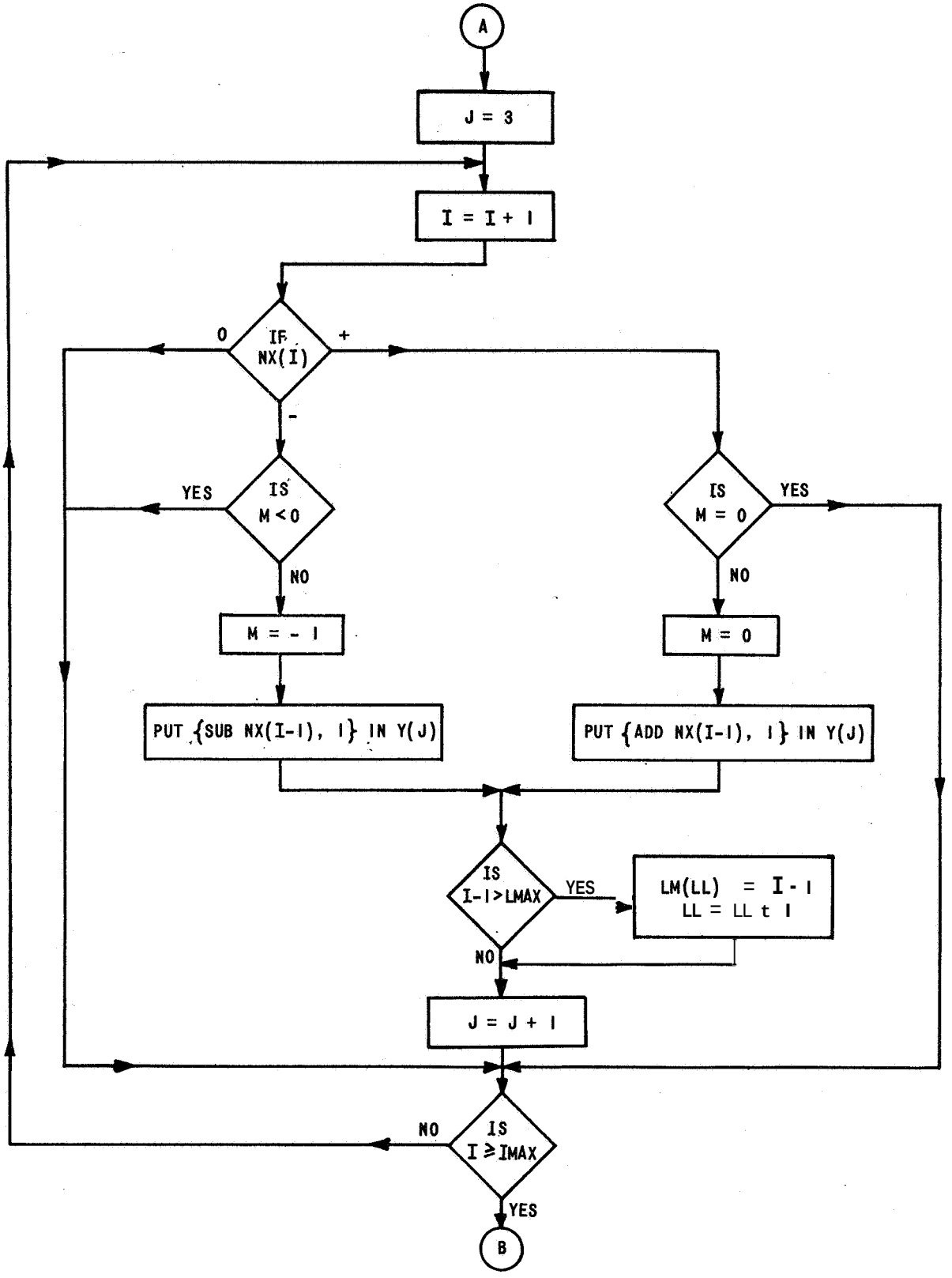


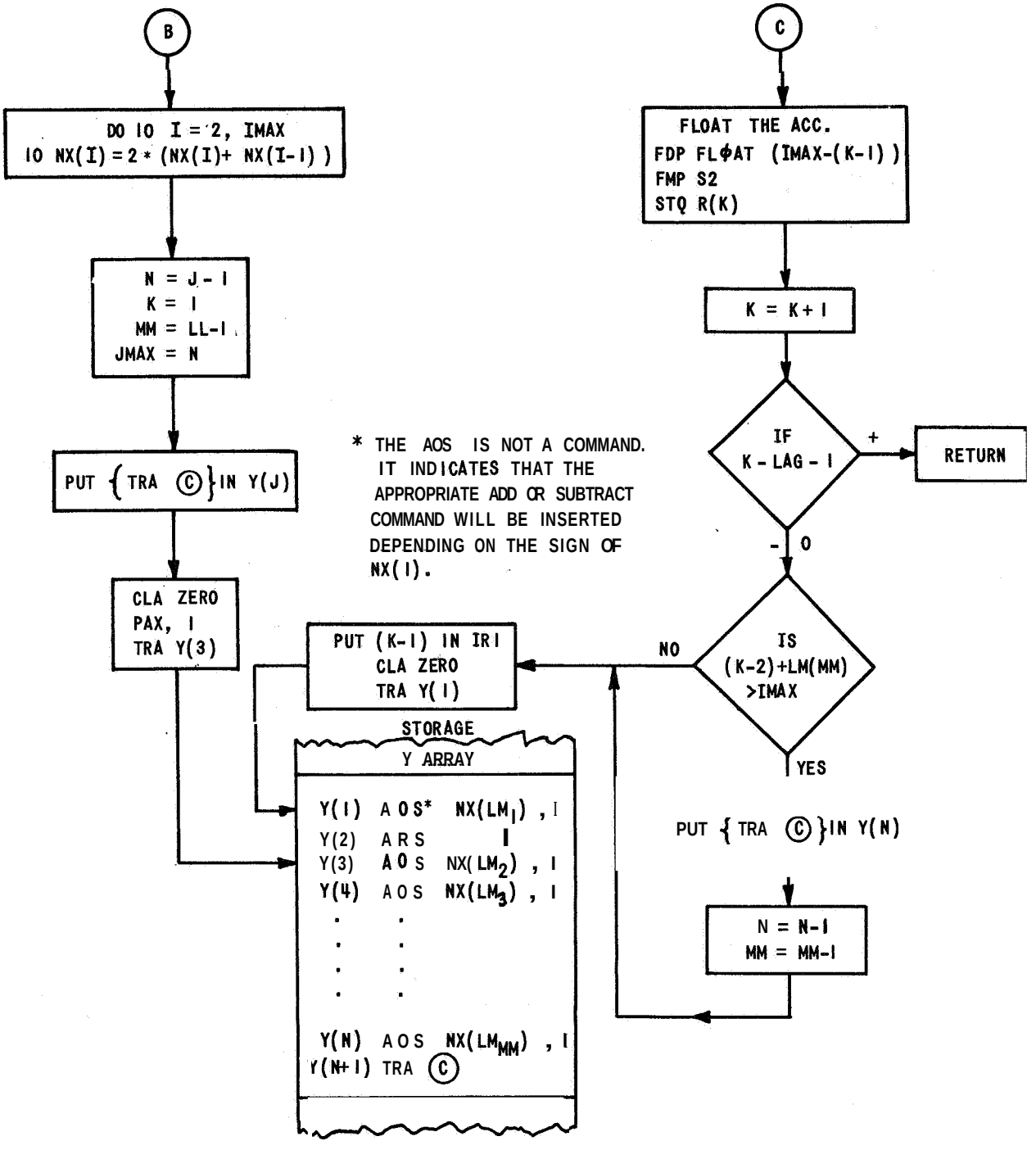
Figure D16 POWER SPECTRUM OF RECORD E12#4 FOR NORMALIZED HALF-POLARITY CORRELATOR

APPENDIX E

Flow Chart of the R_3 correlator.




A
J = 3
I = I + 1
IF NX(I)
0
+
-
IS M < 0
YES
NO
M = - 1
PUT {SUB NX(I-1), 1} IN Y(J)
IS M = 0
YES
NO
M = 0
PUT {ADD NX(I-1), 1} IN Y(J)
IS I-1>LMAX
YES
LM(LL) = I - 1
LL = LL t 1
NO
J = J + 1
IS I ≥ IMAX
NO
YES
B


B
DO 10 I = 2, IMAX
10 NX(I) = 2 * (NX(I)+ NX(I-1))
N = J - 1
K = 1
MM = LL-1
JMAX = N
PUT { TRA © } IN Y(J)
CLA ZERO
PAX, 1
TRA Y(3)
* THE AOS IS NOT A COMMAND. IT INDICATES THAT THE APPROPRIATE ADD OR SUBTRACT COMMAND WILL BE INSERTED DEPENDING ON THE SIGN OF NX(1).
STORAGE
Y ARRAY
Y(1) AOS* NX(LM1) , I
Y(2) ARS 1
Y(3) AOS NX(LM2) , I
Y(4) AOS NX(LM3) , I
Y(N) AOS NX(LMMM) , I
Y(N+1) TRA ©
C
FLOAT THE ACC.
FDP FLΦAT (IMAX-(K-1))
FMP S2
STQ R(K)
K = K+1
IF K - LAG - 1
+
RETURN
-
0
IS (K-2)+LM(MM) >IMAX
NO
YES
PUT (K-1) IN IR1
CLA ZERO
TRA Y(1)
PUT { TRA © } IN Y(N)
N = N-1
MM = MM-1

APPENDIX F
DERIVATION OF THE EXPECTED VALUE OF $R_3'(m)$

$$E\left[R_3'(m)\right] = E\left[\frac{1}{N-m}\sum_{k=1}^{N-m}(f_k + n_k)\,\mathrm{sgn}(f_{k+m} + n_{k+m})\right]$$

$$= \frac{1}{N-m}\sum_{k=1}^{N-m} E\left[f_k\,\mathrm{sgn}(f_{k+m} + n_{k+m})\right]$$

$$+ \frac{1}{N-m}\sum_{k=1}^{N-m} E\left[n_k\,\mathrm{sgn}(f_{k+m} + n_{k+m})\right]$$

$$= \frac{1}{N-m}\sum_{k=1}^{N-m} f_k\, E\left[\mathrm{sgn}(f_{k+m} + n_{k+m})\right]$$

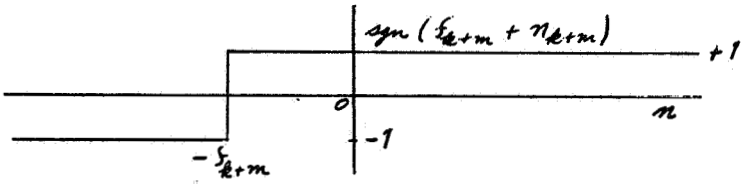
$$+ \frac{1}{N-m}\sum_{k=1}^{N-m} E\left[n_k\,\mathrm{sgn}(f_{k+m} + n_{k+m})\right] \qquad \text{(F-1)}$$

$$E\left[R_3'(m)\right] = A + B$$

Consider A

$$E\left[\mathrm{sgn}(f_{k+m} + n_{k+m})\right] = \int_{-\infty}^{\infty} \mathrm{sgn}(f_{k+m} + n_{k+m})\, f(n;m)\, dn$$

$$A = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \mathrm{sgn}(f_{k+m} + n_{k+m})\, e^{-\frac{n^2}{2\sigma^2}}\, dn$$



Using the above diagram

$$A = \frac{1}{N-m} \sum_{k=1}^{N-m} \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^{f_{k+m}} e^{-\frac{n^2}{2\sigma^2}} dn$$

let

$$\frac{n}{\sigma} = x$$

$$dn = \sigma\, dx$$

$$A = \frac{1}{\pi} \frac{1}{N-m} \sum_{k=1}^{N-m} \int_0^{\frac{f_{k+m}}{\sigma}} e^{-\frac{x^2}{2}} dx \qquad \text{(F-2)}$$

where

$$\sigma = \sqrt{R_n(0)}$$

Now Consider B

$$E\left[n_k \operatorname{sgn} (f_{k+m} + n_{k+m}) \right]$$

let $n_k = u$

$n_{k+m} =$

The joint probability density function of u, v is:

$$f(u, v; m) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} e^{-\frac{u^2 - 2\rho uv + v^2}{2\sigma^2(1-\rho^2)}}$$

where

$$\rho(m) = \frac{R_n(m)}{R_n(0)}$$

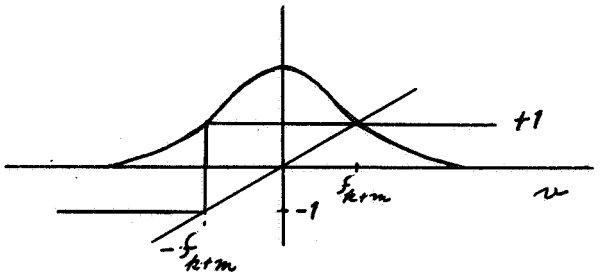
and

$$\sigma^2 = R_n(0)$$

$$E\left[u\,\mathrm{sgn}\left(f_{k+m}+v\right)\right] =$$

$$\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u\,\mathrm{sgn}\left(f_{k+m}+v\right) e^{-\frac{u^2-2\rho uv+v^2}{2\sigma^2(1-\rho^2)}}\,du\,dv$$

$$= \frac{\rho}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} v\,\mathrm{sgn}\left(f_{k+m}v\right) e^{-\frac{v^2}{2\sigma^2}}\,dv$$



$$= \frac{2\rho}{\sqrt{2\pi\sigma^2}}\int_{f_{k+m}}^{\infty} v\, e^{-\frac{v^2}{2\sigma^2}}\,dv$$

let

$$\frac{v}{\sigma} = \beta^{1/2}$$
$$\frac{dv}{\sigma} = \frac{1}{2}\beta^{-1/2}\,d\beta$$
$$\frac{v^2}{\sigma^2} = \beta$$

then

$$E\left[u\,\mathrm{sgn}\left(f_{k+m}+v\right)\right] = \frac{\rho\sigma}{\sqrt{2\pi}}\int_{\frac{f_{k+m}^2}{\sigma^2}}^{\infty} e^{-\frac{\beta}{2}}\,d\beta$$

$$= \sqrt{\frac{2}{\pi}}\,\frac{R_n(m)}{\sqrt{R_n(0)}}\; e^{-\frac{1}{2R_n(0)}f_{k+m}^2}$$

hence the second term B is

$$B = \sqrt{\frac{2}{\pi}} \frac{1}{N-m} \sum_{k=1}^{N-m} \frac{R_n(m)}{\sqrt{R_n(0)}} e^{-\frac{1}{2R_n(0)} f_{k+m}^2} \qquad \text{(F-3)}$$

and

$$E\left[R_3'(m)\right] = \sqrt{\frac{2}{\pi}} \frac{1}{N-m} \sum_{k=1}^{N-m} \int_0^{\frac{f_{k+m}}{\sqrt{R_n(0)}}} e^{-\frac{x^2}{2}} dx$$

$$+ \sqrt{\frac{2}{\pi}} \frac{1}{N-m} \sum_{k=1}^{N-m} \frac{R_n(m)}{\sqrt{R_n(0)}} e^{-\frac{1}{2R_n(0)} f_{k+m}^2} \qquad \text{(F-4)}$$

APPENDIX G
GENERATION OF OPTIMUM SPECTRAL SMOOTHING WEIGHTS

Obtaining an expression for the smoothing weights requires minimizing the expected value of the following index:

$$J = \int_{-\omega_c}^{\omega_c} \left| \Phi_a(\omega) - \Phi_i(\omega) \right|^2 d\omega \tag{G-1}$$

$$\begin{aligned}
\min_{\{D_j\}} E[J] &= \min_{\{D_j\}} E\left[\int_{-\omega_c}^{\omega_c} \left| \Phi_a(\omega) - \Phi_i(\omega) \right|^2 d\omega \right] \\
&= \min_{\{D_j\}} E\left[\omega_s \sum_{j=-\infty}^{\infty} \left(D_j R_j - \phi_j \right)^2 \right] \\
&= \min_{\{D_j\}} E\left[\omega_s \sum_{j=-\infty}^{\infty} \left(D_j^2 R_j^2 - 2 D_j R_j \phi_j + \phi_j^2 \right) \right] \\
&= \min_{\{D_j\}} \omega_s \sum_{j=-\infty}^{\infty} \left(D_j^2 E[R_j^2] - 2 D_j \phi_j E[R_j] + \phi_j^2 \right)
\end{aligned}$$

$$\omega_s \left(2 D_j E[R_j^2] - 2 \phi_j E[R_j] \right) = 0 \qquad j < N$$

$$D_j = \frac{\phi_j E[R_j]}{E[R_j^2]} \qquad j < N \tag{G-2}$$

APPENDIX H
EVALUATION OF THE EXPECTATION OF THE R_2 CORRELATOR SQUARED UNDER GAUSSIAN ASSUMPTIONS

The R_2 correlator is given by

$$R_2(j) = R_j = \frac{1}{N-j} \sum_{i=1}^{N-j} y_i\, y_{i+j}. \tag{H-1}$$

Squaring the expression for R_j in (H-1) yields

$$R_j^2 = \frac{1}{(N-j)^2}\left[\sum_{i=1}^{N-j} (y_i\, y_{i+j})^2 + 2\sum_{k=1}^{N-j}\sum_{\ell=k+1}^{N-j} y_k\, y_{k+j}\, y_\ell\, y_{\ell+j}\right]$$

$$E[R_j^2] = \frac{1}{(N-j)^2}\left[\sum_{i=1}^{N-j} E[y_i\, y_{i+j}\, y_i\, y_{i+j}] + 2\sum_{k=1}^{N-j}\sum_{\ell=k+1}^{N-j} E[y_k\, y_{k+j}\, y_\ell\, y_{\ell+j}]\right]$$

Assuming the process under consideration is Gaussian, the above expression reduces to

$$E[R_j^2] = \frac{1}{(N-j)^2}\left[\sum_{i=1}^{N-j} (\phi_0^2 + 2\phi_j^2) + 2\sum_{k=1}^{N-j}\sum_{\ell=k+1}^{N-j} (\phi_j^2 + \phi_{\ell-k}^2 + \phi_{\ell-k+j}\,\phi_{\ell-k-j})\right]$$

$$= \frac{\phi_0^2 + 2\phi_j^2}{N-j} + \frac{2}{(N-j)^2}\left[\frac{(N-j)^2-(N-j)}{2}\phi_j^2 + \sum_{k=1}^{N-j-1}(N-j-k)\,\phi_k^2 + \sum_{k=1}^{N-j-1}(N-j-k)\,\phi_{k+j}\,\phi_{k-j}\right]$$

$$= \frac{\phi_0^2}{N-j} + \frac{(N-j+1)\,\phi_j^2}{N-j} + \frac{2}{(N-j)^2}\sum_{k=1}^{N-j-1}(N-j-k)(\phi_k^2 + \phi_{k+j}\,\phi_{k-j}) \tag{H-2}$$

Another form of the above expression may be written as

$$E[R_j^2] = \phi_j^2 + \frac{1}{(N-j)^2}\sum_{k=-(N-j-1)}^{N-j-1}(N-j-|k|)(\phi_k^2 + \phi_{k+j}\,\phi_{k-j}) \tag{H-3}$$

APPENDIX I
OTHER REPRESENTATIVE PERFORMANCE MEASURES

1) Choose the measure

$$\theta_1 = \int_0^\infty \dot{u}_\ell^2(\gamma)\,d\gamma \cdot \int_0^\infty \dot{u}_h^2(\lambda)\,d\lambda + \lambda_t^2 \int_0^\infty [1 - u_h(\lambda)]^2\,d\lambda + \lambda_\gamma^2 \int_0^\infty [1 - u_\ell(\gamma)]^2\,d\gamma \qquad \text{(I-1)}$$

where $u_\ell(\gamma)$ is the unit step response of $\ell(\gamma)$ and
$u_h(\lambda)$ is the unit step response of $h(\lambda)$.

This performance measure corresponds to the assumption that the three test inputs are

1) Uncorrelated noise
2) An abrupt step change in the correlation function with time
3) A stationary rectangular correlation function.

The optimal solutions are

$$L(s) = \frac{}{1 + s/\alpha_\gamma} \qquad \alpha_\gamma = \frac{\lambda_\gamma^2}{\int_0^\infty h^2(\lambda)\,d\lambda} \qquad \text{(I-2)}$$

$$H(s) = \frac{}{1 + s/\alpha} \qquad \alpha_t = \frac{\lambda_t^2}{\int_0^\infty \ell^2(\gamma)\,d\gamma} \qquad \text{(I-3)}$$

2) Choose the measure

$$\theta_2 = \int_0^\infty \ddot{r}_\ell^2(\gamma)\,d\gamma \cdot \int_0^\infty \ddot{r}_h^2(\lambda)\,d\lambda + \lambda_t^2 \int_0^\infty \left\{ [\lambda - r_h(\lambda)]^2 + \sigma_t^2 [1 - \dot{r}_h(\lambda)]^2 \right\} d\lambda + \lambda_\gamma^2 \int_0^\infty \left\{ [\gamma - r_\ell(\gamma)]^2 + \sigma_\gamma^2 [1 - \dot{r}_\ell(\gamma)]^2 \right\} d\gamma \qquad \text{(I-4)}$$

where σ_t^2 and σ_γ^2 are arbitrary positive weighting constants.

This performance measure corresponds to the assumption that the three test inputs are

1) Uncorrelated noise
2) An abrupt step changes in the correlation function with time, and a ramp change in the correlation function with time
3) A stationary rectangular correlation function and a stationary triangular correlation function.

Optimal solution A.

$$L(s) = \frac{1 + \left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right)s}{\left(1 + \frac{s}{\rho_1}\right)\left(1 + \frac{s}{\rho_2}\right)} \qquad \text{(I-5)}$$

where

$$\rho_1 = \left|\left(\frac{a + |(a^2 - 4b)^{1/2}|}{2}\right)^{1/2}\right| \qquad \text{(I-6)}$$

and

$$\rho_2 = \left|\left(\frac{a - |(a^2 - 4b)^{1/2}|}{2}\right)^{1/2}\right| \qquad \text{(I-7)}$$

and

$$a = \frac{\lambda_\tau^2 \sigma_\tau^2}{\int_0^\infty h^2(\lambda)\,d\lambda} \quad ; \quad b = \frac{\lambda_\tau^2}{\int_0^\infty h^2(\lambda)\,d\lambda} \qquad \text{(I-8)}$$

for the condition that $a' \geq 4b$

$H(s)$ will have the same form.

Optimal solution B.

$$L(s) = \frac{2(\cos\phi)\rho s + \rho^2}{s^2 + 2(\cos\phi)\rho s + \rho^2} \qquad \text{(I-9)}$$

where

$$\rho = \left|\left(\frac{a + j|(a^2 - 4b)^{1/2}|}{2}\right)^{1/2}\right| \qquad \text{(I-10)}$$

and

$$\phi = \frac{1}{2} \arctan \frac{|(a^2-4b)^{1/2}|}{a} \tag{I-11}$$

for the condition that

$$a^2 < 4b \tag{I-12}$$

$H(s)$ will be of the same form.

In this solution, it is possible for the damping of the complex pole pair to take on values between $0.707 < \zeta < 1.0$, $(\zeta = \cos\phi)$. The parameter, ζ , will, of course, be dependent on a and b.

REFERENCES

1. Anderson, G. W., Aseltine, J. A., Mancini, A. R. and Sarture, C. W., "A Self-Adjusting System for Optimum Dynamic Performance," IRE National Convention Record, Part 4, 1958.

2. Anderson, G. W., Buland, R. N., and Cooper, G. R., "The Aeronutronic Self-Optimizing Automatic Control System," Proc. of Self-Adaptive Flight Control Systems Symposium, WADC Tech. Rept. 54-49, ASTIA AD209389; WPAFB, Ohio, March 1959.

3. Ekre, Helge, "Polarity Coincidence Correlation Detection of a Weak Noise Source," IRE Transactions on Information Theory, Vol. IT 9, No. 1, pp 18-23, January 1963.

4. Tou, Julius T., Digital & Sampled-Data Control Systems, McGraw-Hill Book Co., New York, N.Y., 1959.

5. Farnam, J. J. Jr., Hills, R. Jr., "Correlators for Signal Reception," Technical Memorandum No, 27, NR-384-903, Office of Naval Research, Acoustics Research Lab., Division of Applied Science, Harvard University, Cambridge, Mass., September 1952.

6. Pick, L. A., "A Quasi-Linear Correlator Applied to Signal Location," U. S. Army Signal Research and Development Lab., Fort Monmouth, N. J., June 1961.

7. Schmid, L. P., "Efficient Autocorrelation," Letter to the Editor of Communication of the Association for Computing Machinery, Vol. 8, No. 2, February 1965, pg. 115.

8. Bussgang, J. J., "Cross-correlation Functions of Amplitude Distorted Gaussian Signals," Mass. Inst. Tech. Res. Lab. Electronics, Cambridge, Mass., Tech, Report, No. 216, March 1952.

9. Brown, J. L. Jr., "On a Cross-correlation Property for Stationary Random Processes," IRE Transactions on Information Theory, Vol. IT 3, pp 28-31, March 1957.

10. Leipnik, R., "The Effect of Instantaneous Non-linear Devices on Cross-correlation," IRE Transactions on Information Theory, Vol. IT 4, pp 73-76, June 1958.

11. Madwed, A., "Number Series Method of Solving Linear and Non-linear Differential Equations," MIT Inst. Lab. Report No. 6445-T-26, April, 1950.

12. Tustin, A. , "A Method Analyzing the Behavior of Linear Systems in Terms of Time Series," Journal I.E.E. (Proceedings of the Convention on Automatic Regulators and Servomechanism), Vol. 94, Part 11-A, May, 1947,

13. Anderson, W. H., Ball, R. B., Voss, J. R., "A Numerical Method for Solving Control Differential Equations on Digital Computers," Journal ACM, January, 1960.

14. Blum, M. , "Recursion Formulas for Growing Memory Digital Filters," IRE Trans. Info. Theory, March, 1958.

15. Fryer, W. D., and Schultz, W. C. , "A Survey of Methods for Digital Simulation of Control Systems," CAL Report No. XA-1681-E-1, July 1964.

16. Blackman, R. B., Tukey, J. W., The Measurement of Power Spectra, Dover Publications, 1958.

17. Fleck, John T., Fryer, William D., "An Exploration of Numerical Filtering Techniques, Cornel1 Aeronautical Laboratory Report No. XA-869-P-1, 1 May 1953.

18. Anders, E. B. et al, Digital Filters NASA Technical Report CR-136, December 1964.

19. Laning, H. J. Jr. and Battin, R. H., Random Processes in Automatic Control, McGraw-Hill Book Co., New York, N. Y. s 1956, pp. 269-278.

20. Davenport, W. B. Jr., Johnson, R. A., and Middleton, D. , "Statistical Errors in the Measurements on Random Time Functions," Journal of Applied Physics, American Institute of Physics, New York, N. Y., Vol. 23, April 1952, pp. 377-388.

21. Watts, D. G., "A General Theory of Amplitude Quantization with Applications to Correlation Determination," The Institution of Electrical Engineers, Monograph No. 481M, November 1961.

22. Graham, R. J. , "Determination and Analysis of Numerical Smoothing Weights," NASA Technical Report TR R-179, December 1963.

23. Bendat, J. S., "Principles and Applications of Random Noise Theory," John Wiley & Sons, Inc., 1958.

24. Wierwille, W. W., "A New Approach to the Spectrum Analysis of Nonstationary Signals," IEEE Transactions on Applications and Industry (November 1963), Number 69, pp. 322-327.

25. Fano, R. M., "Short-Time Autocorrelation Functions and Power Spectra," Journal of the Acoustical Society of America, Vol. 22 September 1950, pp. 546-550.

26. Schroeder, M. R. and Atal, B. S., "Generalized Short-Time Power Spectra and Autocorrelation Functions," Journal of the Acoustical Society of America, Vol. 34, November 1962, pp. 1679-1683.

27. Friedman, B., Principles and Techniques of Applied Mathematics, John Wiley and Sons, Inc., New York, 1956, pp. 134-144.

28. Laning, J. H. Jr. and Battin, R. H., Random Processes in Automatic Control, McGraw-Hill Book Co., Inc. New York, 1956, pp. 73-86, 147-165.

29. Benedict, T. R., and Bordner G. W., "Synthesis of an Optimal Set of Radar Track-While-Scan Smoothing Equations," IRE Transactions on Automatic Control, Vol. AC-7, July 1962, pp. 27-32.

30. Wierwille, W. W., "Experimental Study of a New Method of Time Delay for Analog Computers," (To appear in the IEEE Trans. on Electronic Computers, Vol. EC-14, August 1965).

31. Gardner, Harvey L., "The Residue Number System," IRE Transactions on Electronic Computers, June 1959, pp. 140-147.

32. Cheney, Philip W., "A Digital Correlator Based on the Residue Number System," IRE Transactions on Electronic Computers, March 1961, pp. 63-70.

33. Szabo, Nicholas, "Sign Detection in Nonredundant Residue Systems," IRE Transactions on Electronic Computers, August 1962, pp. 494-500.

34. Kein, Y. A., Cheney, P. W., and Tannenbaum, M., "Division and Overflow Detection in Residue Number Systems," IRE Transactions on Electronic Computers, August 1962, pp. 501-507.

35. Wierwille, W. W., "A Theory of Nonstationary Correlation Analysis" appearing in CAL Project Spectrum Final Report, CAL No. UA-1889-B-1, Contract NAS 1-3485, 15 January 1965.